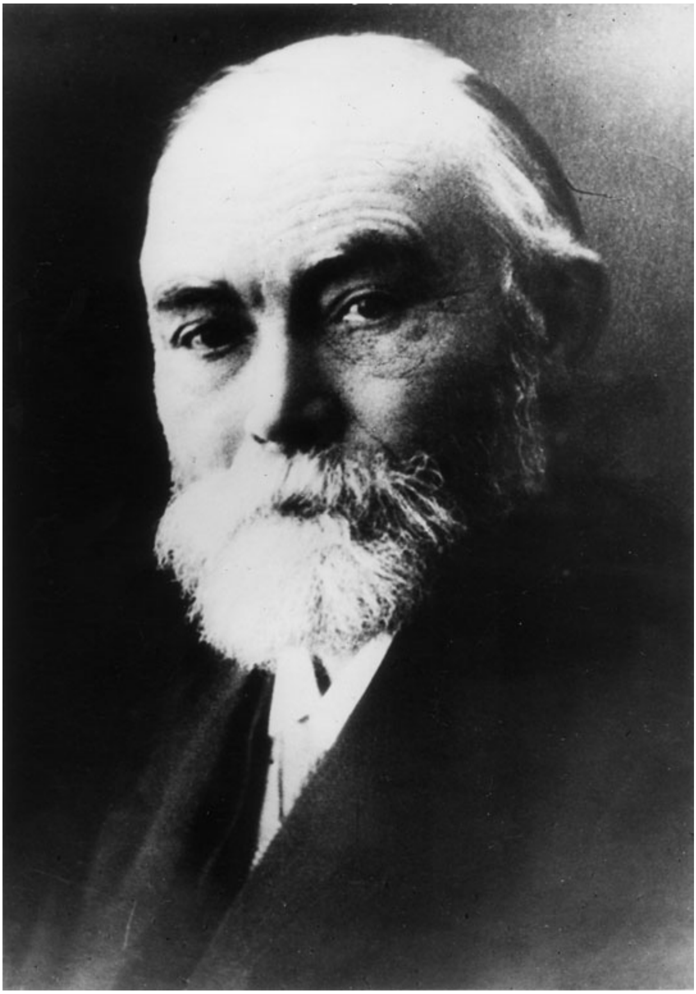
TYLER BURGE

truth thought reason

ESSAYS ON FREGE

Truth, Thought, Reason



Gottlob Frege, c. 1920. Photograph © AKG London.

Truth, Thought, Reason

Essays on Frege

TYLER BURGE

CLARENDON PRESS · OXFORD

OXFORD
UNIVERSITY PRESS

Great Clarendon Street, Oxford OX2 6DP

Oxford University Press is a department of the University of Oxford.
It furthers the University's objective of excellence in research, scholarship,
and education by publishing worldwide in

Oxford New York

Auckland Cape Town Dar es Salaam Hong Kong Karachi Kuala Lumpur
Madrid Melbourne Mexico City Nairobi New Delhi Shanghai Taipei Toronto

With offices in

Argentina Austria Brazil Chile Czech Republic France Greece
Guatemala Hungary Italy Japan South Korea Poland Portugal
Singapore Switzerland Thailand Turkey Ukraine Vietnam

Published in the United States
by Oxford University Press Inc., New York

First published 2005

British Library Cataloguing in Publication Data
Data available

Library of Congress Cataloging in Publication Data
Data available

ISBN 0–19–927853–9
ISBN 0–19–927854–7 (pbk.)

10 9 8 7 6 5 4 3 2 1

Typeset by Kolam Information Services Pvt. Ltd, Pondicherry, India
Printed in Great Britain
on acid-free paper by
Biddles Ltd., King's Lynn, Norfolk

Preface

I read Frege in graduate school, of course. But I became seriously engaged only when I began teaching his work in the 1970s at UCLA. Each year that I gave the course, I spent three-quarters of the time on *The Foundations of Arithmetic*. On my own, I devoted many hours to formalizing his definitions, doing derivations with them, and making handouts for students. *The Foundations of Arithmetic* engendered enthusiasm, both in my attitude toward the teaching and in the students' response. Frege's powerful criticisms of alternative views made a deep intellectual impression, and his obvious struggles in developing his own positive positions on fundamental philosophical issues were inspiring. Frege's philosophy of mathematics was the center of the course. His papers in the philosophy of language came last, occupying only a quarter of the time and providing "glimpses beyond". I centered on the philosophy of mathematics because I thought it had to be understood, because UCLA's quarter system made courses relatively short, and because Frege's philosophy of language was taught in other courses. An unforeseen by-product of this approach was that it helped me see how Frege's epistemology—his attempt to understand the nature of mathematical knowledge—lies at the philosophical heart of all his work, including his philosophy of language.

From the beginning, I found Frege's epistemology and his association of language with thought attractive. I was attracted to his rationalism. I found his concentration on thought and knowledge as expressed in language a welcome alternative to relatively narrow reflection on linguistic structures, which had been my starting point in philosophy. My developing philosophical commitments in the theory of reference and philosophy of mind made me sensitive to central points that Frege neglected or made mistakes on. But I thought that many of his positions were right and profound. I knew that there was much to be gained from reflecting on his views on language, thought, and knowledge. The first two papers that I published on Frege (both in 1979) tried to work within Frege's point of view as much as possible. Yet neither paper employed as strict a historical methodology as I later came to believe was necessary for the most effective understanding and presentation of Frege's views.

At the time, I regarded both of these papers as "holiday" work—work to be done with the left hand, so to speak, as a diversion from my main work in philosophy. I valued thinking along with Frege as I would value going through a mathematical proof or a physical workout. It seemed to be worthwhile in making one clearer-headed and stronger. And it was fun. In those first years, I did not see it as *directly* enhancing my own systematic work in philosophy.

I had had some training in history when I was in college. In coming up against ways in which Frege was philosophically foreign, I came to recognize that Frege should be studied as a figure in the history of philosophy. With the paper "Frege on Extensions of Concepts: 1884–1903" (1984), I tried to apply historical methodology in a more rigorous way.

While I was writing this paper, an event occurred that marked what I think of as a more substantial change in my approach to the study of Frege and to the history of philosophy. In one of the last revisions of the paper, before submitting it for publication, I came suddenly to see that Frege's rationalism guided his conception of sense in a much deeper way than I, or perhaps anyone, had appreciated before. I came to recognize how profoundly different Frege's conception of sense is from modern conceptions of communal linguistic meaning. The differences between modern views and Frege's views about demonstratives and proper names—which I and others had previously recognized—came to seem only the tip of a very large, strange, and wonderful iceberg. Some of these differences are discussed in the last sections of that 1984 paper, in "Frege on Sense and Linguistic Meaning" (1986), implicitly in "Intellectual Norms and Foundations of Mind" (1986),[1] and in subsequent papers on Frege. What I want to emphasize here, however, is not the content of my realization (the way that Frege's rationalism guided his conception of sense): I want to emphasize the effect that the realization had on me.

I retain an absolutely vivid image of the moment when the realization came to me. I was working alone late at night in a warmly, but dimly lit dining room—not my usual place for working on philosophy. The idea seemed to erupt like a sharp explosion and then to spread like a lava flow. I was completely absorbed. I was keyed-up, but in a way that did not affect a concentration that lent some weight, in retrospect, to the old metaphors of the intellect's being emancipated from time. It was the sort of moment of insight and discovery that one is granted only occasionally, but which sustains intellectual life—both in the thrill of the initial revelation and in the gradual realization and working out of consequences and connections. It is the intellectual counterpart of falling in love in a way that stays solid, develops, and deepens beyond the initial excitement. The experience was a heightened instance of how—for all the frustration of struggling through the difficulty of the subject, and all the tediousness of being careful to put things in order

[1] *The Journal of Philosophy*, 83 (1986), 697–720.

and avoid foreseeable mistakes—philosophy can be both thrilling and life-sustaining.

The experience changed my attitude toward engagement with Frege, and indeed toward the role of the history of philosophy in my broader philosophical work. Studying Frege became not merely an exercise in intellectual hygiene and development. It became a means of philosophical discovery. There is much that remains to be understood about Frege. There are whole reaches of his thought that offer the possibility of breakthrough insights to students of his work. I invite the reader to join in the quest for further discovery.

In landing at UCLA after ejecting from graduate school, I was very fortunate to have three senior colleagues who had an exceptionally deep understanding of Frege. I had several helpful conversations with Montgomery Furth when I was starting out. Furth's translation and introduction to *The Basic Laws of Arithmetic* and his "Two Types of Denotation" were invaluable in my early Frege education.[2] Deplorably, the former work is out of print. Furth's writing about Frege remains among the best introductions.

Alonzo Church was a sometimes ghostly, sometimes substantial, presence in my early years at UCLA. Some of us used to joke that Church was Frege himself, having learned English—with his German origins thinly disguised by the overlay of a Virginia accent—and having picked up more than a few inches and pounds. Church was not a person one conversed idly with. I audited his courses, read his work, and was influenced by his intellectual standards and by the power of his pragmatic rationalist point of view.[3] Often, experiencing the presence of a great intellectual has an effect that goes well beyond what mere reading or even listening could achieve. My experience of Church was of this sort. Church did as much as anyone to make Frege's work effective in twentieth-century philosophy. His particular impact on my understanding of Frege was very substantial—in inspiration, in the formation of standards, and in developing instincts for Frege's ways of thinking. Furth and Church are no longer among the living. So my thanks to them must be correspondingly attenuated.

The third colleague who influenced my understanding of Frege remains a colleague. I have greatly benefitted from the intellectual example and historical-philosophical instincts of David Kaplan. Kaplan had been a student of both Church and Carnap. His early work was in formal Fregean semantics. He has maintained a pedagogical and systematic interest in Frege throughout his career. Although Kaplan disclaims using historical methodology, he reads as

[2] Gottlob Frege, *The Basic Laws of Arithmetic: Exposition of the System*, translated and edited, with an introduction by Montgomery Furth (Berkeley, University of California Press, 1967); Montgomery Furth, "Two Types of Denotation", in *Studies in Logical Theory*, American Philosophical Quarterly Monograph Series, Monograph no. 2 (Oxford, Basil Blackwell, 1968).

[3] I came to be an editor of Church's work, a task I began in the late 1970s: *The Collected Works of Alonzo Church* (Cambridge, Mass.; MIT Press, forthcoming).

closely as anyone I know. His historical instincts about Frege are reliable and true. Repeatedly, I have found, by checking texts, that Kaplan's "off the cuff" claims about Frege go to the heart of both large structure and subtle nuance in Frege's conceptions.

As any teacher must acknowledge, I have benefited from the enthusiasm, questions, and insights of many students. Among these, I specially think of Nathan Salmon, Marco Ruffino, and Simon Evnine all of whom went on to write on Frege.

I have also learned a great deal from Charles Parsons and Tony Anderson over the years. Both have been valuable interlocutors. The occasions on which we interacted are few in number, but large in impact. More recently, I have been stimulated in fruitful ways by discussions of Frege with Christopher Peacocke.

Finally, I want to acknowledge more personal debts. My family supported and endured—chiefly my wife Dorli, but also my sons Johannes and Daniel and my parents Mary (now deceased) and Dan.

Contents

Acknowledgments

"Frege", in H. Burkhardt and B. Smith (eds.), *Handbook of Ontology and Metaphysics* (Munich: Philosophia Verlag, 1991).
"The Concept of Truth in Frege's Program", *Philosophia Naturalis*, 21 (1984), 507–12.
"Frege on Truth", in L. Haaparanta and J. Hintikka (eds.), *Frege Synthesized* (Dordrecht: Reidel, 1986), 97–154.
"Frege and the Hierarchy", *Synthese*, 40 (1979), 265–81. Copyright © by D. Reidel Publishing Co., Dordrecht and Boston.
"Sinning Against Frege", *Philosophical Review*, 88 (1979), 398–432.
"Frege on Sense and Linguistic Meaning", in D. Bell and N. Cooper (eds.), *The Analytic Tradition* (Oxford: Blackwell, 1990), 30–60.
"Frege on Extensions of Concepts, From 1884 to 1903", *The Philosophical Review*, 93 (1984), 3–34.
"Frege on Knowing the Third Realm", *Mind*, 101 (1992), 633–49.
"Frege on Knowing the Foundation", *Mind*, 107 (1998), 305–47.
"Frege on Apriority", in P. Boghossian and C. Peacocke (eds.), *New Essays on the A Priori* (Oxford: Oxford University Press, 2000), 11–42.

Abbreviations

AA	Kant, *Gesammelte Schriften*
B	Frege, *Begriffsschrift und andere Aufsätze*, ed. I. Angelelli
BL	Frege, *The Basic Laws of Arithmetic*, ed. and trans. M. Furth
'C & O'	Frege, "Concept and Object"
CP	Frege, *Collected Papers on Mathematics, Logic, and Philosophy*, ed. B. McGuinness
'F & C'	Frege, "Function and Concept"
'F & O'	Frege, "Function and Object"
FA	Frege, *Foundations of Arithmetic*, trans. J. L. Austin
FG	Frege, *On the Foundations of Geometry and Formal Theories of Arithmetic*, ed. E. W. Kluge
FPL	Michael Dummett, *Frege: Philosophy of Language*
G & B	*Translations from the Philosophical Writings of Gottlob Frege*, eds. Peter Geach and Max Black
GG	Frege, *Grundgesetze der Arithmetik*
IFP	Michael Dummett, *The Interpretation of Frege's Philosophy*
Kl	E. D. Klemke, ed., *Essays on Frege*
KS	Frege, *Kleine Schriften*, ed. I. Angelelli
LI	Frege, *Logical Investigations*, ed. P. Geach
N	Frege, "Negation"
NS	Frege, *Nachgelassene Schriften*, eds. H. Hermes, F. Kambartel, and F. Kaulbach
PM	Bertrand Russell and A. N. Whitehead, *Principia Mathematica*
PMC	Frege, *Philosophical and Mathematical Correspondence*, eds. B. McGuinness and H. Kaal

PW	Frege, *Posthumous Writings*, trans. P. Long and R. White
'S & R'	Frege, "On Sense and Reference"
TOE	Michael Dummett, *Truth and Other Enigmas*
WB	Frege, *Wissenschaftlicher Briefwechsel*, eds. G. Gabriel, H. Hermes, F. Kambartel, C. Thiel, and A. Veraart

O stands for 'original'.

The abbreviations used in individual chapters are cited again within the chapters, to make each chapter self-contained.

Introduction

Gottlob Frege (1848–1925) is surely the great philosopher least known to the general intellectual public in proportion to his stature and effect on philosophy. He was relatively unknown and unappreciated even by mathematicians and philosophers during his lifetime. Toward the end of his career, his contributions were promoted by Russell, Wittgenstein, and Carnap—and slightly later by Church. He came to be recognized as the fountainhead of mainstream philosophy in the twentieth century—the century's most important and influential philosopher. His main work was published between 1879 and 1903. It came to life and had its effect only subsequently.

There are many currents in twentieth-century philosophy. What I called the "mainstream" is the approach to philosophy that has dominated the English-speaking world for most of the last century. It has been central in Poland and Scandinavia, and is becoming increasingly prominent in the rest of continental Europe. In fact, it is showing every sign of becoming global. This approach draws more of the best minds and carries on a more coherent and progressive discussion than any other approach within philosophy. It is this approach that owes more to Frege than to anyone else. For much of the twentieth century, this mainstream was called "analytic philosophy". I shall return to this label.

Although Frege is well appreciated in philosophy, his name and work remain unknown to the wider intellectual public. The reason for this obscurity is not hard to come by. Frege's writing is narrowly focused in philosophy of mathematics and philosophy of language. Some of it is technical. The subjects themselves are forbidding. Yet most of the passages of Frege's writing that have had philosophical impact are not specially technical, and are not hard to read, as philosophy goes. Students new to his work warm to it quickly, because of its crystalline clarity, the bold symmetries of its structure, and its appeal to intuition and common sense in discussion of abstract topics like the nature of number or the relation between thought and communication.

I mentioned that Frege's contributions are a source of what has often been called "analytic philosophy". The term is popularly associated with an emphasis on clarifying meaning and centering philosophical discussion on language and logic rather than on "reality". It is associated in the popular mind with expelling from philosophy terms or problems not fit for scientific or rigorous discussion, and with a negative attitude toward traditional philosophical methods and issues. It is commonly thought that this approach is too narrow to be of general interest. Hostility to philosophy, so conceived, is widespread not only in popular culture, but among the general intellectual public.

This popular view once had considerable merit. It still has some truth in it. Among the first appropriators of Frege's work were the logical positivists who dominated philosophy in the period between the 1920s and 1950. Although among Frege's three original promoters, Russell and Wittgenstein have greater individual philosophical stature than Carnap, Carnap was part of a movement—logical positivism—that had the broadest influence among those who first made use of Frege's work. Logical positivism has long been overthrown as a doctrine. Yet it cast a large shadow into the latter half of the twentieth century. Most of the charges of the preceding paragraph have some force against this movement and at least some of its successors. The charges are, however, one-sided and constitute a caricature, even of logical positivism.

The charges are one-sided inasmuch as they miss the genuine methodological progress that occurred in philosophy through the first half of the century. Owing much to Frege's development of modern logic, his alliance of philosophy with scientific attitudes and norms, and his clarity in exposition and rigor in argument, the standards for philosophical discussion took a giant leap forward. The effect of this progress was to make philosophy a much more communal enterprise. Philosophical theses could be discussed more fruitfully. Reasons and criticism could be more easily assessed. Grounded agreement or disagreement became more common. Even though philosophy is not, in general, a science, it took on some of the best methodological characteristics of science.

The charges constitute a caricature inasmuch as they fail to recognize the seriousness of the basic problems that engaged the positivists. The positivists inherited from Frege an interest in meaning as expressed in language. Both Frege and his positivist successors were concerned to understand language and meaning, because they saw such understanding as a new and promising route to understanding the nature of human knowledge. Frege and the positivists took human knowledge to be best exemplified by scientific knowledge. (As we shall see, however, they understood this point in very different ways.) This approach to philosophy through a consideration of the nature of human knowledge is, of course, traditional. It had dominated the subject from Descartes through Kant. Kant was a source of inspiration for both Frege and his positivist successors.

In broadest terms, Frege added two things to this tradition. He added a concentration on language in the expression of knowledge. And he added a recognition of the power of logic to illuminate the structure of language and its contribution to the expression of knowledge. Frege is responsible for establishing and developing modern logic—a huge achievement. He used logic to brilliant effect in investigating the structure of language and the ways it connects to the world and expresses thought. He used this investigation to better understand human knowledge. In this respect, Frege should be seen as continuing and making major contributions to a central tradition in philosophy.

Frege's positivist successors used Frege's innovations. Yet they did so against a background of philosophical attitudes that Frege did not share. They differed with him, in absolutely fundamental ways, about both meaning and knowledge.

Although the positivists took up Frege's focus on language, they appended to it an ideology of exclusivism. The only meaning that might be of any cognitive value was, for them, scientific meaning. It is this ideology that led positivism to be aggressively deflationary about those aspects of philosophy and culture that it could not assimilate to a scientific paradigm. Frege was specially interested in the language of the mathematical sciences. But his doctrine exhibits no such exclusivism. It exhibits no inclination to deflate philosophical problems that are not at the center of his scientific focus.

Whereas Frege's successors saw themselves as overturning traditional philosophy, Frege saw himself as continuing a philosophical tradition. Whereas his positivist successors applied reductionistic or deflationary attitudes to nearly all philosophical problems, Frege confined his reductionism to the attempt to reduce the mathematics of number to logic. He shows no special inclination to hold that reduction is a good method in philosophy generally. He shows no inclination to hold that the task of philosophy is to show that philosophical problems amount to less than meets the eye.

The positivists began with the hypothesis that there is something cognitively defective about any meaning that is not the meaning of scientific language. They began by demanding an explanation of meaning that can be distilled down to scientific elements. Notoriously, they tried to reduce meaning to procedure for verification. It is, however, the method of their approach that I want to highlight. The notion of meaning was assumed from the beginning to need reduction, deflation, reshaping, or explanation in terms that were "scientifically acceptable". This attitude has continued in philosophy well after the demise of positivism. In fact, forms of scientism, reductionism, and deflationism have remained prominent in philosophy since the positivist initiative.[1] Frege is certainly concerned with ways in which

[1] The deflationist approach marks the work of Wittgenstein and his successors. Wittgenstein was one of Frege's earliest champions, as noted; but he was never strictly a positivist. So his motives for deflating philosophical problems are different—less centered on science. The same deflationary approach occurs in the work of Quine, who was the most prominent force in the overthrow of positivism, but who in this respect was an ally of Carnap and the positivists. Quine's work has played a large part in keeping alive the positivist demand for a theory of meaning and the positivist view that philosophical methods and problems that are not methods and problems within current science are to be doubted or dispensed with.

ordinary language can be an obstacle to the progress of science. But his more open or pragmatic approach to philosophy stands clear of these trends set in motion by the positivists. I believe that his approach will remain valuable as its scientistic-reductionist-deflationist offshoots diminish in prominence.

Frege postulated two types of "meaning": (a) *reference* or *denotation* (*Bedeutung*), a relation between language and its subject matters, and (b) *sense* (*Sinn*), the way of thinking relevant to the truth or falsity of the thought that is directly associated by the individual language-user with his linguistic expressions.[2] Frege provided deep and lasting accounts of the structures of these types of meaning. He showed no interest in explaining them in terms that are more scientific according to some preconceived standard. Nor did he think that they have to be ignored or dispensed with in a final science. I believe that he would have taken senses to be part of the subject matter of a genuinely public, scientific psychology and linguistics as these disciplines are now being pursued. I see no textual basis for holding that he thought that the semantics of a scientific language is not every bit as factual as any other special science. His work on the two types of meaning was the historical basis for all major subsequent work in semantics—in philosophy, mathematical logic, and linguistics.

Frege's positivist successors differed with him fundamentally about knowledge as well as about meaning. Frege's fundamental motivation, like that of the positivists, was to understand human knowledge.[3] Whereas

[2] The translation of Frege's term "*Bedeutung*" has been a source of controversy. There is something to be said for leaving the term untranslated. But with few exceptions, I do not follow this path. In the early translations, the German term was rendered variously as "reference", "nominatum", "designatum", "denotation". There followed a period in which it was translated as "meaning". This latter translation is the ordinary rendering of the German "*Bedeutung*". But it is deeply misleading as regards Frege's usage. Frege's term is very clearly a technical term. And it is a term that does not correspond well at all to the more usual, commonsense understanding of the term "meaning". In its simplest applications it corresponds fairly well to the ordinary term "reference". I believe that Frege did partly motivate his use of his technical term "*Bedeutung*" by considering simple cases of reference. Like "reference", and unlike the more usual construals of "meaning", "*Bedeutung*" indicates a word–subject–matter relation. But "reference" tends to make Frege's view that sentences and function signs have a *Bedeutung* unnecessarily odd-sounding. And (like "meaning") it does not prepare the reader for how theory-laden Frege's term "*Bedeutung*" ultimately is. "Nominatum" is probably worse than "reference" as a translation, in that it suggests even more strongly that "*Bedeutung*" is purely about naming or singular reference. It is not. Carnap used the translation "designatum". I think that this is a better translation than "nominatum", even though it too suggests, though less strongly, singular reference. My choice not to use it rests primarily on stylistic and historical considerations. I prefer to avoid the heavy-handed Latinate form. And I want to suggest a relation between Frege's usage and a certain broad historical tradition. Church used "denotation" as translation. I have followed his usage at least in this introduction and in my later work on Frege. The translation risks too close an association with Russell's famous use of the same term, which use is, of course, rather different. But it is a relatively technical-sounding English term—an advantage. And it connects Frege's usage to a broad tradition—which existed independently of Russell—in which two types of "meaning" were distinguished: denotation and connotation, extension and intension. Frege's particular version of a two-"meaning" distinction is, in my view, not just specific to his theory, but far superior to previous versions of such a distinction.

[3] I believe that Russell's philosophy is also fundamentally driven by epistemic concerns. His approach to the theory of knowledge is closer to Frege's than to the positivists'. Russell was a

positivism—in fact, most post-Fregean mainstream philosophy until the last decade or so—was aggressively empiricist, Frege was a rationalist. That is, positivism was centrally motivated by the view that all genuine knowledge depends for the force of its warrant or justification on sense experience. Frege believed that knowledge in logic and arithmetic depends for justification only on reason.

Logical positivism rests on two doctrines: (a) that statements in mathematics and logic are *analytic* in the sense that their truth does not depend in any way on relations to a subject matter, but depends purely on their meaning; and (b) that the cognitive meaning of a statement consists in its method of verification. (Roughly, cognitive meaning is meaning that can be rationally construed as involved, or purportedly involved, in a true or false statement.) The second doctrine ties cognitive meaning to science. The first maintains that logic and mathematics do not constitute genuine knowledge or genuine science: knowledge of a subject matter is restricted to the natural sciences. Both doctrines were overthrown, I think decisively, in the middle of the twentieth century.[4]

The positivists took their paradigm science to be physics. Frege's paradigm science was mathematics. Frege thought of proof in mathematics, ultimately in logic, as the paradigm type of justification or "verification", although of course he recognized inductive justification in empirical science. The positivists took proof in mathematics to be merely an instrument within natural science with no independent status. Proof in pure mathematics did not count for them as justification of any genuine knowledge at all. They took experiment in natural science to be the paradigmatic justification.

These differences regarding justification and knowledge are associated with differences about ontology—regarding what exists or what "has being". The positivists took mathematics and logic to be true independently of a subject matter. This is the first of the two doctrines mentioned above. The positivists helped produce a philosophical climate in which suspicion of abstract entities—whether numbers, functions, or senses—was axiomatic for many philosophers. Frege exemplifies the natural ontological attitude of working mathematicians, a relaxed ontological Platonism. *Ontological Platonism* is the view that certain abstract entities exist or have being, and their

rationalist and Platonist about mathematics, like Frege. Like Frege he regarded logic as a general science of "being". In my view, however, his basic doctrines about knowledge are extremely crude in comparison to Frege's. I think that Frege's conception of sense (as thought component and as idealized cognitive value) is on the right track for understanding cognition. I believe that Russell's account of denotation is of little value as the primary theoretical notion in understanding cognition. What I call Frege's pragmatic rationalism is vastly more sophisticated than Russell's theory of acquaintance.

4 The key papers were W.V. Quine, "Two Dogmas of Empiricism" (1953), repr. in *From a Logical Point of View* (New York: Harper and Row, 1961); and "Carnap and Logical Truth" (1954), in *Ways of Paradox* (New York, Random House, 1966). For discussion of them, see my "Philosophy of Language and Philosophy of Mind: 1950–1990", *The Philosophical Review*, 101 (1992), 3–51, and "Logic and Analyticity", *Grazer Philosophische Studien*, 66 (2003), 199–249.

being and natures are independent of relations to any entities that exist, or have being, in time. Frege took numbers and functions to be a subject matter of mathematics and saw no scientific reason to hold that they occur in space or time, or depend for their being or natures on anything that occurs in space or time. He extended this attitude to his account of sense or thought content. I believe that Frege's Platonism is a source of both philosophical insight and philosophical excess. I shall return to this matter.

A further difference between Frege and his most influential successors lies in their views about logic. Positivism was committed to seeing logic, like mathematics, as lacking a subject matter. Often this commitment took the form of seeing logic as a formal tool for aiding inference in the empirical sciences, but as having no "substantive" content of its own. Frege saw logic as a scientific language expressing knowledge of a subject matter, in this respect like any other science. He regarded its meaning or content as just as substantive as that of any other scientific language. He took its truths to be just as dependent on relation to a subject matter as the truths of any other science. In fact, he saw logic as the most general science of "being". Here too Frege is part of the main tradition in the philosophy of logic. I believe that he used this point of view to make fundamental contributions to understanding logical truth and logical consequence.

I mentioned that Frege is the source of what is commonly called "analytic philosophy". I indicated a preference for the term "mainstream twentieth-century philosophy". "Analytic philosophy" is a term quite appropriate to the work of Frege's positivist successors, but it is at best misleading as applied to Frege's own work. It has become a commonplace within philosophy to note that the term "analytic philosophy" is now used so loosely, and has come to apply to such a diverse discipline, that there is no general substantive characterization of it. Only historical and extremely broad (almost uninformative) methodological characterizations seem to fit. "Analytic philosophy" can be appropriately construed as a proper name, not a description. Yet in its early applications the phrase had solid descriptive aspects. It suggested a central concern with meaning cut off from metaphysics or "reality". It suggested an approach to philosophy through a method of analysis. It suggested the prominence of analytic truths among philosophical results. These suggestions remain in the general intellectual consciousness.

I have already criticized the popular idea that philosophy during this period is primarily concerned with linguistic meaning. Frege, Russell, and the positivists were concerned with linguistic meaning *because* they saw it as the key to understanding human knowledge.

It is true that parts of philosophy of language and philosophy of logic became relatively autonomous in the latter part of the twentieth century. Language and meaning came to be studied for their own sake. Parts of philosophy are now continuous with semantics in linguistics. Parts are now continuous with semantics in mathematical logic. These are healthy philosophical developments with ample precedent in the history of philosophy. Philosophy has always been midwife to new sciences. Frege deserves the largest credit for both of these developments, even though they occurred after he had done his main work.

But even now, the main line of philosophical development continues to link understanding language with a larger range of philosophical topics. These topics include the way that language expresses knowledge, the relation between meaning and communication, the transmission of knowledge through communication, the way that understanding meaning aids in understanding mind, the way that understanding particular discourses contributes to understanding the topics of those discourses—from aesthetics and ethics to scientific explanation. In these traditional philosophical enterprises, Frege's work on language has been a stepping stone to understanding the whole range of philosophical problems, including traditional philosophical problems.

The popular picture that concern with linguistic meaning in analytic philosophy proceeds independently of concern with "reality" is a version of the mistake embodied in the positivist doctrine that logic and mathematics are analytic. This is a mistake that Frege did not make. His conception of sense and his semantical doctrines are inextricable from his concern with the nature of "reality" or being. Here again Frege's revolutionary innovations are in the service of a type of philosophical inquiry as old as Plato and Aristotle.

What of Frege and "analytic method"? There is a very broad sense in which Frege's method is analytic. He seeks to isolate basic concepts and basic principles in trying to demonstrate logicism—the view that the mathematics of number is reducible to logic. This enterprise of seeking basic concepts and principles is again common to many of his philosophical forebears—Plato, Aristotle, Descartes, Leibniz, Kant.

It would be a fundamental mistake, however, to see Frege as carrying out this method by analyzing definitions of words through sheer reflection on the words or through focusing entirely on common linguistic usage—even usage within science. In fact, Frege's way of trying to understand linguistic meaning, linguistic structure, and fundamental principles is notably synthetic. It is fundamental to his attempt to understand meaning that his investigation proceed by considering the whole range of inferences that are expressed in the use of any particular piece of language. He sought understanding through holistic, systematic reflection on the discursive uses of language.

Frege does not assume that existing usage is determinative of what language expresses. He allows any element of simplification, and any piece of rational insight or discovery about the subject matter, to be potentially relevant to understanding the nature and structure of senses and denotations—his two types of meaning. Reflection on whole systems of inferential and other cognitive activity is fundamentally a synthetic enterprise. Relevant systems are not treated as given (there to be simply analyzed) independently of ongoing inquiry. So conceiving Frege's method as analytic is at best deeply misleading.

The third descriptive connotation associated with the term "analytic philosophy" is that philosophy so named is concerned with analytic truths. Frege does have a concept of analytic truth. But he explains it simply as truth provable from general laws of logic. Analytic truths are, definitionally, just a subclass of logical truths. The notion is associated with a certain form of justification—proof from logical laws. Relative to the issues associated with analytic truth in the twentieth century, his concept of analytic truth is philosophically non-committal.[5]

The sense of "analytic truth" that the positivists urged and which entered into the popular conception of analytic philosophy is different from Frege's. That notion of analytic truth is expressed in the first of the two positivist doctrines. It is the idea of a truth made true by meaning alone, independently of any subject matter. It is the notion employed by the positivists to cut off mathematics and logic from epistemic and metaphysical inquiry. The positivists regarded empiricism as a fundamental philosophical commitment. Empiricism claims that all genuine knowledge of any subject matter is warranted through sense experience. If logic and mathematics were true independently of any subject matter, they could not threaten empiricism. And they could not engender metaphysical questions about the nature of mathematical entities. Many positivists held that where philosophy makes true statements, as opposed to practical recommendations, those statements are also analytically true—in the sense that their truth does not depend on a subject matter.

Frege's philosophy recognizes no analytic truths in this sense. His primary interest lies in the nature of human knowledge of mathematics. He engages in deep reflection about the nature of the objects and functions that this knowledge is about. He takes the knowledge that he gains through reflection on meaning (knowledge of sense as well as denotation) to illuminate the nature of human thought, and indeed the nature of all knowledge. So his vision of philosophy could not be further from the picture advanced by the positivists. They pictured philosophy as simply producing practical recommendations and analytic truths, which delimit its own ambitions and scope. Frege's philosophy is a philosophy of inquiry and discovery.

I have emphasized some fundamental ways in which Frege differs from the most influential appropriators of his work, and from the popular conception of the approach to philosophy that he helped initiate. One might think that these

[5] He does call the relevant truths of logic "analytic" because he regarded such truths as "contained" within the basic logical laws. Traditionally, containment was taken to be a type of analyticity. This idea embodies a mathematical error. It is a mathematical error because by Gödel's incompleteness theorems, in higher-order logics, including second-order logic, not all logical truths are contained in the basic logical truths in the sense that Frege intended. Frege counted such higher-order logics as part of logic. And he understood containment in terms of logical derivability.

are aspects of Frege's work that simply got left behind. I believe that this is not true. The matter is, however, complex.

The dissolution of the distinctive spirit of positivism has been a much slower process than the collapse of its letter. That spirit is still alive. But it is not dominant. It contends with eclectic, pragmatic approaches to philosophy driven more by the diversity and difficulty of philosophical problems than by deflationary ideologies.

Conceptions of meaning, particularly those associated with the problems that Frege introduced his notion of sense to solve, are still in flux in the wake of the failure of the verificationist theory of meaning. There are approaches not moved by the positivist spirit. One such approach is that of taking semantical notions roughly at face value and using them as refined theoretical primitive notions in a semantical theory. This was Frege's method.

The implications of the failure of the positivist conception of analytic truth have not been fully digested. But the appeal to analyticity, in the positivists' sense, is no longer prominent in epistemic and ontological enterprises. Moreover, Frege's rationalist conception of knowledge of mathematics and logic has undergone a notable revival. An openness to inquiry, which his philosophy exhibits, has begun to replace the predominantly reductionist and deflationist spirit that marked much of the philosophy influenced by his work. I believe that the aspects of Frege's philosophy that I have emphasized will prove to be longer-lived than contrasting counterparts. I believe that these aspects will be a continuing source of inspiration in philosophy.

Frege's philosophy should be of intellectual interest to the broader intellectual public, in any case. He set an example for philosophical reasoning and method that had the largest effect in producing a very deep change in the standards for philosophical discussion. This change has seeped into every area of philosophy, not just the areas that Frege and his immediate successors concentrated upon. Its effect has been to make a larger portion of serious philosophy accessible to a broad intellectual public than perhaps at any other period since the period of the Greeks.

Of course, the most original, serious philosophy is never easy. The subject matter is abstract. Progress often requires precision and a precisely understood vocabulary. One simply has to read serious philosophy more slowly than many other subjects. But much of Frege's writing lies in the tradition of Plato, Descartes, Hume, Russell—presenting difficult matters in a style that can be grasped by intelligent non-specialists. Frege's example changed standards for clarity of expression and argumentative rigor.

Historical interest in his work is justified by a reason that should motivate all historical interest: It is a means to better self-understanding and better realization of one's own pursuits—in this case philosophical and, more generally, intellectual pursuits.

Then, of course, Frege's philosophy is intrinsically interesting. It is old enough to be fruitfully foreign. Yet it has a currency and vitality that belie its age. The essays collected here explore aspects of Frege's philosophy that are of current and, I believe, long-term philosophical interest.

Before providing a more specific introduction to the philosophical orientation of this collection of essays, I want to discuss my motive and method in writing them.

My motive has always been philosophical. Frege has an obvious and special relevance to philosophy in our time. One way to gain a perspective that distinguishes the deep from the superficial is to reflect on great thinkers of the past. Frege had profound and lasting insights that are fundamental for the theory of knowledge, philosophy of language, philosophy of mind, philosophy of logic, and philosophy of mathematics. Understanding a great philosopher in depth helps deepen one's own thinking. Some of the most important philosophical values of studying Frege emerged for me only after years of study and reflection. I hope to communicate some of these values.

My method in writing about Frege has been, almost from the beginning (see the Preface), steadfastly historical. History of philosophy is a branch of philosophy as well as a branch of history. Approaching history as part of an ongoing philosophical enterprise, however, carries dangers that are more acute than in other types of history. In studying the great figures of the past in the service of truth in philosophy, it is easy to assimilate their ideas to more recent ones or to make mistakes in criticizing them because one overlooks differences in their terminology, aims, or background assumptions. These tendencies are especially easy to fall into if one is too quick to argue with them or treat them as contemporary interlocutors, opponents, or allies. To appreciate the full depth of a great philosopher of the past, one must be patient. One must maintain a reflective listening attitude that goes beyond what is needed to understand a contemporary.

There is, first, temporal distance to compensate for. Presuppositions change. What seems obvious changes. Reader-expectations change. Familiar terms' meanings or at least connotations change. One can master these changes only through systematic reading and rereading, and through the consistent exercise of historical as well as philosophical judgment and perspective.

There is, further, the depth of genius to accommodate to. Appreciating the size, richness, detail, and originality of the conceptions of a great philosopher again requires patience in listening, in reflecting, and in willingness to read and reread in systematic and comparative ways.

Given the dangers of parochialism, of mis-measuring historical distance, and of underestimating philosophical genius, I have found it important to support my exposition through substantial citations of texts. This can seem tedious and overly scholarly. It is, however, an important control. Attempting to "tell a story" that is not tied down at every turn to textual evidence commonly becomes an expression of the story-teller's own ideology and commonly fails to appreciate the complexity, depth, and foreignness of a great philosopher's work.

Of course, given my philosophical motivation, historical exposition is guided by my philosophical interests. Historical method can help counteract tendencies toward glib disagreement and ideologically based "stories" based on incomplete understanding. But it has its own dangers. The expenditure of energy that goes into getting a figure historically right can sap the historian's philosophical energy. Historical work can fall into slavish agreement or uncritical reporting. I hope to have avoided these pitfalls. Although I present Frege's thinking without constant evaluation, I have definite views about what is right and what is wrong in Frege. (I should add that these have changed in some ways with deeper understanding of his work.) Most of my differences with Frege, and the deeper, more subtle ways in which Frege has been a positive influence on me are, I think, best left to venues in which I present my own philosophical work.

Historical perspective is, however, served by evaluative comment. Since I see my reflections on Frege as an adjunct to my own philosophical enterprises, I cannot resist connecting the two at least at some of the most central junctures. I hope that these essays will be read in the light of, and as part of, my own philosophical contributions. Still, in these essays I have tried to train the historical focus on Frege. Whether I have supplemented this focus with philosophical judgment in an illuminating but unobtrusive way is for each reader to decide.

The essays collected here do not in any sense comprise a complete view of Frege's philosophy. There are many aspects of his work that I do not touch. Some are mentioned as background or drawn on but taken for granted. Many have been well discussed by others. The essays center on issues in Frege's work that are of special philosophical interest to me.

In some cases the essays are more difficult to read than Frege's own work. Understanding them well benefits from having at least some background in Frege. In discussing pros and cons of interpretation and offering substantive evaluation, they are sometimes more complex than at least the surface of Frege's own writing. Still, I hope that patient, slow reading, will find them forthright and clear. I hope that in combination with more direct reflection on Frege, they will yield philosophical insight and stimulation in the reader.

The essays center on Frege's views about "meaning" and knowledge—the two topics cited earlier on which Frege differs most deeply from the logical positivists. This neat categorization is made more complex by the fact that Frege conceives of "meaning" as being of two types (*Bedeutung* and *Sinn*). It is also made more complex by the fact that his investigation of "meaning" is driven by a certain conception of truth as an aim that underlies language and science, and by the fact that his investigation of truth and "meaning" are intertwined with his views about knowledge.

After a brief overview paper, "Frege" (1991), the essays are divided, somewhat awkwardly, into three parts. Part I concerns Frege's investigation of structure—the structure of language and thought, which he sees as integral to the structure of knowledge. Frege's work on truth as the aim of logic and science and his account of the structure of "meaning" are the seeds from which all his philosophical contributions grow. I am interested in Frege's investigation of structure because it exhibits his *method* of coming to understand linguistic and epistemic structure. An offshoot of this investigation is his account of the structure of sense. Indeed, the very conception of sense as idealized cognitive value is understood ultimately in terms of its role in aiding pursuit of truth. Part II deals primarily with Frege's conception of sense as thought, or thought component, and as idealized cognitive value. Part III centers on Frege's rationalist conception of knowledge. Obviously, the three parts should not be regarded as compartmentalized. In Frege's own work, the issues are interwoven. So reflection on the topics of any one part should deepen one's understanding of the topics of the others.

I would like now to make some more specific philosophical remarks about Frege's contributions, as discussed in these three parts. These remarks are meant to serve primarily as *philosophical* orientation on the largely historical accounts that the essays offer.

PART I: TRUTH, STRUCTURE, AND METHOD

The project that drove Frege's life work was an attempt to establish logicism. Logicism is the view that the mathematics of number can be reduced to logic. Establishing logicism requires defining the primitive mathematical expressions or notions in terms of expressions or notions of pure logic, and then using the principles of logic together with the definitions to prove the axioms and theorems of the relevant mathematics. Frege hoped to show that the mathematics of number—principally arithmetic and analysis—could be thus reduced to logic. He regarded geometry as a different type of mathematics, not reducible to logic.

Frege conceived his logicist project as a contribution to the theory of knowledge. He wanted to explain the fundamental character of our knowledge of arithmetic and analysis. Although this philosophical motivation is very clear, Frege proposed to carry out his project with the rigor of a project in mathematics. In doing so, he established standards for rigor unprecedented in philosophy and hardly precedented in mathematics.

Frege realized that to carry out the logicist project he needed a precise logical system. To this end, he developed modern first-order and higher-order logic. This was by far the largest step forward in logic since Aristotle. In fact, to oversimplify for effect without fundamentally distorting the historical or philosophical facts, this was the first major step forward in logic since Aristotle. With it Frege completed first-order and second-order logic proper—logic considered independently of its meta-theory.[6] Frege's development of logic is one of the great achievements in intellectual history.

Frege also realized that to carry out the logicist project, he needed to understand the logical structure of principles of mathematics and logic in a sufficiently definite way to be able to carry out the relevant proofs. This realization led Frege to three of his deepest insights. These insights can be usefully seen as part of the following idealized reasoning (though I do not claim that Frege used precisely this reasoning). Thoughts or principles are most perspicuously expressed in sentences. Reflectively understanding the logical form or logical structure of sentences requires understanding the logical structure of their component parts. Logical structure is revealed in the way that good deductive inference hinges on structure. So understanding logical structure of sentences and their component parts depends on understanding the structure of good deductive inference. Understanding the structure of deductive inference depends on systematic reflection. For logical inference can combine *any* sentences or thoughts in a single argument, and by looking at numerous combinations one recognizes structural patterns that otherwise would not be salient. So to understand logical structure of sentences and their component parts, one must reflect on the way that sentences enter into a wide range of inferential combinations. The point of inference is to preserve truth in making transitions from true premises to true conclusions. So in reflecting on linguistic structure as it is revealed in inference, one should focus on the contributions of elements in such structure to determining the truth of sentences and to preserving the truth of sentences in inference. So to understand reflectively the structure of sentential parts, one should reflect systematically on their contribution to conditions under which sentences count as true, and thus on their contribution to determining conditions under which truth is preserved in deductive inferences.

[6] Frege gives an account of higher-order logic as well, but all the power essential to higher-order logic is, in effect, contained in second-order logic. Scholarship in the history of logic has, in the last few decades, emphasized numerous significant achievements in logic between Aristotle and Frege. But, seen from the largest perspective, the picture that I have sketched still seems to me correct. The sketch relies on philosophical as well as historical oversimplification. It is a philosophical issue exactly what to count as logic. For some, Russellian type theory with an axiom of infinity counts as logic. For others, classical set theory counts as logic. On these conceptions, Frege's contributions to logic stop well short of completing logic proper (logic considered independently of its meta-theory). Frege's own logic included an axiom that turned out to be contradictory, and its role in his theory has been filled by axioms in type theory or set theory. As I read the situation, however, the most nearly standard conceptions of logic take logical axioms not to be committed to an infinity of entities. These conceptions omit not only Frege's contradictory axiom but the axioms that were proposed as substitutes for it in type theory and set theory. So first-order and higher-order logics as they are now standardly conceived are less than what Frege proposed as logic. But they are all indebted to the work of Frege. Although Frege made serious contributions to the meta-theory of logic, much of what is now standard meta-theory (starting, for example, with the completeness theorem for first-order logic) was produced after Frege's career was completed.

The three key insights are present in the preceding sequence of reasoning:

(1) Thought is fruitfully understood by reflecting on language.[7]
(2) One understands the structural nature of thoughts and components of thoughts not by taking those natures for granted, nor by relying on simple intuition about grammar or thinking, nor by invoking general philosophical *dicta*, but by reflecting discursively on a large number of deductive inferences among sentences.
(3) The key to understanding such structure lies in understanding the contribution of such components to determining truth conditions and to preserving truth in deductive inference.

These three insights came to dominate twentieth-century reflection on language and thought. They have proved to be fruitful and genuine. They have had their effect in tandem. The first insight led to the fruitful work in philosophy that I mentioned earlier. It also led to explosive developments in linguistics, psycholinguistics, and cognitive psychology. But it would not have had its effect if it had not been linked with the second and third insights.

The second and third insights had already been enunciated by Kant, at least in germ.[8] Kant held that concepts are essentially predicates of judgment. He held that judgment aims at truth and is essentially propositional. Thus he regarded components of thoughts as having their function and structure only in the context of propositional judgment and inferences among judgments.

It is likely that Frege's insights gained something from his exposure to Kant. But Frege's insights in this domain went deeper than anything one finds in Kant, or anyone else before Frege. Frege applied the second and third insights with a rigor and system that were simply unprecedented. Frege had a full, precise logic to work with. Through this means he gained a grip on the nature of numerous logical inferences that no one before him had. Frege applied the method of investigating the sub-structure of thoughts and sentences in a systematic and detailed way. This application made the insights come to life in accounts of the contributions of specific structures to specific inferences. Frege also combined the second and third insights with the first. He centered his reflection on thought in an investigation of language. This

[7] Michael Dummett takes analytic philosophy to be defined by two doctrines: that a philosophical account of thought can be obtained through a philosophical account of language, and that a comprehensive account of thought can only be obtained through a philosophical account of language (*Origins of Analytical Philosophy* (Cambridge, Mass.: Harvard University Press, 1994), 4). I believe that this characterization of analytic philosophy is interesting but much too narrow. It excludes Frege and Russell from being analytic philosophers, because they reject the second doctrine. There are many others in the analytic tradition who would follow them in this rejection. Dummett's characterization fails to apply to yet others in the tradition who are silent on one or both of the doctrines. Dummett has taken views held by many analytic philosophers to characterize a tradition that is better characterized historically, and by reference to a loose sharing of methods and approaches. These do indeed center, to an unprecedented degree, on language. But attitudes toward language within the movement vary. What has been widely known as analytic philosophy is not a philosophical ideology.

[8] Kant, *Critique of Pure Reason*, A65–9/B90–4.

provided reflection with a concreteness of application not easily obtained by other means.

In *The Foundations of Arithmetic* (1884) Frege enunciates an idea that is closely associated with the second of the three insights. This enunciation is what is known as his "context principle". In fact, writing as if he were stating a single principle, Frege presents or implies, at different places in the book, three non-equivalent formulations. These are, I believe, correctly seen as three different principles.[9] When one takes account of Frege's later (1891) distinction between sense and denotation, each of these three formulations divides into two further non-equivalent formulations. In the end, I think it correct to say that Frege believed in at least six "context principles". Although Frege's formulations are terse, and although his application of the principles leaves room for interpretation, I shall state what I take the principles to be.

The first two principles are methodological:

Always seek to understand the denotation of a word not in isolation but in the context of its role in a sentence.

Always seek to understand the sense of a word not in isolation but in the context of its role in a sentence.

The second pair of principles concern necessary conditions:

A word relates to a denotation only through its having a role in a sentence or in sentences.

A word relates to a sense only through its having a role in a sentence or in sentences.

The third pair concerns sufficient conditions:

If a word is a component (under structural analysis) in a true sentence, it has a denotation; and the denotation is fixed through its contribution to the sentence's being true and to the preservation of truth in good deductive inference.

If a word is a component (under structural analysis) in a true sentence, the word has a sense; and the sense is fixed through its contribution to the cognitive value and its role in contributing to the cognitive aspects of good deductive inference.

The first pair simply states the method of investigation that I characterized earlier. The second pair provides a philosophical rationale for the method. The third pair—particularly the principle that concerns denotation—provides a philosophical basis for understanding theoretical reference to abstract entities and other theoretical entities of science. I would make certain qualifi-

[9] The formulations occur in *The Foundations of Arithmetic*, Introduction, p. x, sections 60, 62, 106. In some cases Frege states a necessary condition but uses it as a (qualified) sufficient condition.

cations on the principle before accepting it. But I think that it is deeply insightful inasmuch as it locates ontology in the evaluation of theories. It implicitly rejects philosophical requirements on reference or denotation that go beyond standard methods for evaluating the truth of sentences or theories. For example, the principle implicitly rejects the requirement that to make reference to an object through thinking about it, one must have a mental image of the object, or bear a causal relation to the object. It rejects such requirements unless they are motivated as requirements on the evaluation of the truth of sentences or on the evaluation of theories containing the sentences. I believe that all six principles are, with relatively minor qualifications, sound.

The third of the three insights is associated with Frege's recognition that judgment (or belief) and assertion—which aim at truth—are key to understanding thought and language. Frege's logical work is motivated in a remarkably single-minded way by his attempt to understand principles governing truth. He sees logic as codifying such principles because logic attempts to understand preservation of truth in argument. The centrality of these motivations is the topic of both the short paper "The Concept of Truth in Frege's Program" (1984) (Chapter 2 below) and section I of "Frege on Truth" (1986) (Chapter 3 below). The 1984 paper is essentially extracted from the 1986 paper. I include it because I think it brings into focus two major instances in which the depth of Frege's concentration on truth in motivating his views has been underestimated. The longer paper deals with these same issues, but I believe that isolating them in the shorter paper gives them more impact.

One instance is the attribution of the Church–Gödel argument that all true sentences denote the same thing to Frege. This elegant and fascinating argument was certainly inspired by Frege's reasoning. But the attribution of it to him misses how central his conception of the centrality of truth in logic is. The other instance is Dummett's well-known criticism of Frege's treatment of sentences as denoting truth-values. This criticism underestimates how carefully Frege distinguishes sentences—vehicles for assertion—from names in his logic. I discuss this matter in more detail below.

Frege's three key insights led to further insights. They led first to a working conception of truth-conditional semantics. As mentioned earlier, Frege's practice, and many of the details of his work, have remained mainstays in semantics worked out in philosophy, linguistics, and mathematical logic.

Frege was probably the first to emphasize, in the context of a systematic study of language, the creativity of language use. He noticed that everyone uses new sentences every day. In effect, we are in principle competent to use and understand an infinity of sentences. He realized that this was possible only because the sentences are built out of a finite number of building blocks and a finite number of formal principles of construction. This insight reappeared in Chomsky's work and was applied in understanding the grammar of natural language. Frege was centrally interested in the logical form and semantics of a language suited to science. These basic insights, and the example of his practice of finding structure through comparing numerous sentences and their relations in inference, underlie the main development of the scientific study of all language. This study has deep implications for understanding the nature of thinking.

Frege's analysis of structure in terms of its contribution to specifying conditions under which a sentence is true led to his distinction between propositional content (the bearer of truth or falsity) and force (the attitude or speech act that makes use of the content). For example, Frege drew a clear distinction between the logical operator *negation* and the speech act *denial*, and between a thought content and the judgment (or assertion or supposition) of the thought content. These distinctions have been a basis for more detailed theoretical distinctions between semantics and pragmatics.

The focus on truth conditions helped Frege draw his famous distinction between sense and denotation. By reflecting clearly on what conditions in the world make sentences or statements true, he realized that the cognitive value associated with component expressions must differ from their denotations or references. Frege associates his notion of sense essentially with determination of truth conditions. No alternative conception of sense or meaning has been any stronger or more fruitful.

Frege's notion of sense has been controversial. As I will explain below, much of the controversy rests on misunderstanding. But aspects of his use of the notion are certainly mistaken. Frege makes commitments in his theory of thought, in which his notion of sense is a central feature, that ultimately cannot be sustained. Nevertheless, the explanatory *role* that Frege gave his notion of sense—that of representing cognitive value—must, I think, be filled by some substantially similar theoretical notion. Frege's notion of cognitive value is exceptionally idealized. But his basic insights can be used to account both for actual human cognition and for a more idealized potential for cognition. I believe that Frege himself used his notions of cognitive value and sense in these ways.[10]

Frege associates his notion of sense, partly but essentially, with determination of truth conditions. I am broadly sympathetic with the development of a conception of meaning or sense that is relevant to determining truth conditions. I believe that no alternative conception of sense or meaning has been stronger or more fruitful.

Frege used the focus on truth conditions to help distinguish between what is part of the cognitively relevant "meaning" of a sentence (both denotation and sense) and what is part of the associated tone, or other more loosely related cognitive or linguistic associations with the sentence. This distinction was taken up by most other subsequent theorists of language.

[10] For more on the extreme degree of Frege's idealization of cognitive value, see the last part of my discussion of sense in Part II of this Introduction, in the last section of my Postscript to "Frege on Truth", and in "Frege on Sense and Linguistic Meaning" (Ch. 6 below).

Frege's second and third insights not only opened fundamental distinctions in semantics and in the epistemology of language use. They helped him make two important contributions to the meta-understanding of logic.

One is a contribution to the understanding of the fundamental notions of meta-logic—logical validity and logical consequence. Frege provided a non-modal semantical explication of the way logical truth and deductive logical consequence depend on the form of sentences and their component parts. Frege's insights were taken up by Tarski and Skolem to provide the first mathematically rigorous explication of the traditional non-modal notion of formal consequence (and logical truth explicable in terms of formal logical structure). Frege's explications are not as systematic as Tarski's. But they are clearly in a meta-logical tradition that explicates logical truth and good deductive inference in terms of form and structure rather than in modal or epistemic terms. Frege is a central figure in a tradition that includes Abaelard, Scotus, and Bolzano, and which construes logical truth and logical consequence in these intuitive terms.

Most representatives of this tradition hold that logical truths and good deductive inferences are necessary and can be known in ways that are justified independently of sense experience. The explication of logical truth and good deductive inference in this tradition relies, however, on logical form and logical structure, rather than on necessity or some analog of apriority. The notion and role of explication vary from one representative of this tradition to another. Frege regarded semantical explication as non-fundamental from an epistemic point of view. But he took semantics seriously. In fact, he was the first author to develop semantics in a systematic way—the way that eventually flowered in its full mathematicization. I see no reason to think that he doubted that semantics yields knowledge.

In any case, it seems clear that Frege not only belongs to the tradition of understanding logical truth and logical consequence in formal or structural terms. He made fundamental contributions to that tradition—given his clarity about logical form and semantical structure, his association of logical form and structure with maximal generality (rather than modality), and his association of semantical structure with truth.[11]

The other fundamental contribution to meta-logical understanding is a contribution to untangling the traditional problem of predication. This problem is that of explaining the difference between predication and the relation between a predicate and what it predicates. Plato pointed out that if in a sentence like "Theaetetus sits", the name "Theaetetus" stands for the man, and "sits" stands for the property (or form) of sitting, there is a problem of explaining how the sentence differs from a list in which expressions standing

[11] Cf. "Frege on Apriority" (Ch. 10 below) and Postscript to "Frege on Truth". For much more extensive discussion of these matters, see my "Logic and Analyticity". An Appendix to "Logic and Analyticity" contains a short history of the role of the intuitive notion of (formal) logical consequence in the history of logic.

for the man and the property of sitting occur in succession. Plato observed that introducing a third entity—participation or inherence—does not by itself solve the problem if it becomes simply one more entity to be stood for. (So the list *Theatetus, inherence, sitting* gets us no further.) Plato makes the obvious point that the verb "sits" makes a different grammatical contribution to a sentence from any contribution made by its nominalization, "sitting". As far as I know, Plato did not take the issue much further. Understanding the peculiar nature of verbs (or, more generally, predicates) in a systematic way remained a problem that continued to worry philosophers through the centuries.[12]

It would seem to follow from the last point which I attributed to Plato that even if "sits" and "sitting" relate semantically to the same entity, and even if the relation is the same, the grammatical standpoint from which they relate to the entity must be different. Even if they both denote the same entity, what they do within the sentence with respect to the entity is significantly different. That is to say, whether or not "sits" and "sitting" (or "the property of sitting") stand for entities, and whether or not they stand for the same entity, their roles in the sentence in connecting the relevant entity (if any) to other entities are different. Whether these different roles bear different semantical relations to the subject matter, or whether (on the contrary) they bear the same relations but constitute grammatically different relata, seems to me a secondary issue. Some have thought that one must hold that one or both of such words do not stand for any entity at all. Some have held that they stand for different entities. I believe that none of these positions in themselves illuminates the original problem. The problem lies not at the level of the entities stood for, but in the grammatical or logical roles of the expressions (or thought components) within sentences or thoughts.

The problem is to explain the role of predication in such a way as to distinguish it from whatever semantical relation a predicate bears to the world. The predicate bears whatever relation it bears to the world regardless of whether it is predicated of something. "Sits" bears its semantical relations —stands for the same property or is true of the same objects—regardless of what sentence it occurs in, or even whether it occurs in a sentence. When "sits" comes into combination with "Theaetetus" in a sentence, it does something more: it attributes a property of (or to) an individual, or otherwise constitutes predication. How is this "something more" to be understood in a systematic way?

[12] Donald Davidson, *Truth and Predication* (Cambridge, Mass.: Harvard University Press, 2004). Davidson thinks that the key to unraveling the difficulty lies partly in denying that predicates are semantically related to anything, like properties, besides the objects that they are true of. I believe that this diagnosis is incorrect. The key to solving the problem lies at the logical/grammatical level, not at the ontological level. This is, in effect, the insight that Church completed, building on Frege. For more on this, see n. 13.

The question is so close to bedrock that it is hard to know what sort of explanation would be illuminating. But Frege found a way to make a fundamental contribution to answering this question. He accepted the point that the semantical relation associated with predicates differs from the semantical relation associated with names, though he saw both as types of denotation—as types of relation between an expression and a subject matter.

Frege then mobilized two ideas. First, he explained the semantical relation associated with predicates in terms of the relation between a functional expression and a function. This idea enabled him to explain predication in terms of functional application. A function word stands for the same function at all times; but when it is grammatically combined with a name (or other expression for an input or argument of the function), the combination stands for functional application and yields the value of the function. Thus Frege explained predication in terms of functional application. The distinction between merely standing for something and predicating that something of something else was illuminated through a mathematical operation that has firm and independent explanatory power.

Second, Frege associated predication with its use in judgment and assertion, consequently with their objective or aim—truth. The point of judgment and assertion is to arrive at or present the truth. Frege associates those functional applications which for him constitute predications with the functional value, truth. So the application of the function that "sits" functionally stands for to the man that "Theaetetus" stands for yields truth as the value of the functional application. Of course, some predications fail their objective. Such predications yield falsity. Frege illuminated predication by associating it with a mathematical operation integrated into a semantical account that illuminates the point of predication, and linguistic use more generally.

Although Frege is correctly credited with deep insight here—an elaboration of the second and third insights that I discussed earlier—he is commonly seen as having made two serious mistakes. He seems to have thought that in view of the deep difference between the semantical relations associated with predicates and names, the two cannot bear their relations to the same sort of entities. Thus he would hold that "sit" and "sitting" (or "the property of sitting") cannot stand for the same entity. It is sometimes also suggested that he thought that if names and predicates were to bear semantical relations to the same entities, one would have to treat sentences as lists of names or lists of predications.

These would be fallacious inferences. It certainly does not follow from the deep differences between names and predicates as regards their grammatical roles that a name cannot stand for the same entity (a function) that a function expression functionally stands for. A difficulty in Frege's position did emerge when he committed himself to asserting the German analog of the sentence "The concept (or function) *horse* is not a concept (or function)". Since the subject term "The concept *horse*" has the grammatical role of a name, Frege thought that it could not stand for a concept (or function). This is deeply counterintuitive. I believe that this constitutes one of Frege's most serious mistakes. Church showed, in his calculus of lambda-conversion, that there is no reason why a syntactically unsaturated expression—one with no open argument places or free variables—cannot denote a function, including a concept. The key point is to distinguish their grammatical roles.[13]

The second common criticism of Frege is that in taking predicates to stand for functions whose values are truth or falsity, he assimilated sentences to names—names of the values or outputs of predicational functions. It is commonly pointed out that it is artificial to take truth and falsity to be objects, and even more strange to take sentences to be, in effect, names of those objects. Over many years, Michael Dummett held that Frege's views here do constitute an assimilation of sentences to singular terms. He insisted that Frege's assimilation constitutes a fundamental error. He maintained that the assimilation is incompatible with Frege's own insight, elaborated in the context principles, that sentences are deeply different from components of sentences in their use and place in the language. In expressing thoughts that are true, or at least thoughts that are used to aim at truth, sentences are central to judgment, assertion, and inference in ways that singular terms are not.[14] Perhaps because they accord with a sense of the unnaturalness of taking sentences to denote truth-values as objects, Dummett's criticisms seem to be widely accepted and often repeated.

I believe that I have shown in "The Concept of Truth in Frege's Program" (1984) (Chapter 2 below) and (with more contextual background) in sections II and III of "Frege on Truth" (1986) (Chapter 3 below) that Dummett's criticisms are thoroughly mistaken. It is true that names and sentences are semantically similar in Frege's theory, in that they both denote objects. However, Frege maintains, even within his logic, a significant distinction between sentences and names. What Dummett thinks of as sentences are in

[13] Alonzo Church, "The Calculi of Lambda-Conversion" (1941), in T. Burge *et al.* (eds.), *The Collected Words of Alonzo Church* (Cambridge, Mass.: MIT Press, forthcoming). Church himself took the result of attaching the lambda operator to an expression in order to produce a further expression with no free variables as denoting a function. He took the corresponding functional expression with free variables as denoting its values "ambiguously". I leave open whether this latter "take" is the best way to regard the situation. The important point, I believe, is that saturated expressions—expressions which are syntactically without open argument places or free variables, and which cannot, as they stand, function syntactically to make attributions or to take arguments—can denote functions or concepts. Church separates the syntactic role of being a predicate (or, more generally, a functional expression) from the role of relating semantically to an entity that can be attributed (or, more generally, to a function). I discuss philosophical grounds as to why Frege may have reasoned himself into his mistaken position in "Frege on Extensions of Concepts, From 1884 to 1903" (1984) (Ch. 7 below). For an illuminating analysis of Frege's mistake, which I largely accept—with qualifications deriving from the points made above—see Terence Parsons, "Why Frege Should Not Have Said 'The Concept *Horse* is not a Concept' ", *History of Philosophy Quarterly*, 3 (1986), 449–465.

[14] Michael Dummett, *Frege: Philosophy of Language*, 2nd edn. (Cambridge, Mass.: Harvard University Press, 1981), 7, 184, 196. Cf. also Dummett's paper, "Truth" (1959), repr. in *Truth and Other Enigmas* (Cambridge, Mass.: Harvard University Press, 1979).

Frege's actual logic *treated* as nominalizations of sentences. In fact, when Frege glosses unasserted names of truth-values like (like "2 + 3 = 5") in *The Basic Laws of Arithmetic*, he uses nominalizations like "2 + 3's being equal to 5" or "the truth-value thereof that if something is the square root of one then its not being the fourth root of one". (See, for example, sections 5, 8, 12, 13 of *The Basic Laws of Arithmetic*.) Thus what Dummett thinks of as "Snow is white" is understood in Frege's theory as "snow's being white". "Snow's being white" does name a truth-value, and "white" does designate what Frege calls a "first-level concept"—a function from individuals to truth-values. But in Frege's theory the real occurrence of sentence-forming predication, an operation crucial to sentences' being assertable, lies in the occurrences of the horizontal sign, which translates as "is the true". It is a deeply significant fact, which Dummett completely overlooks, that the only expressions that are judgeable in Frege's theory (i.e. the only ones to which the vertical sign, the judgment stroke, can be applied) are expressions that begin with the horizontal sign.

So only expressions like "Snow's being white is the true" are assertable—not expressions that we would normally take to be formalized as sentences. Thus Frege takes "W(s)" to be a nominalization. All genuine *sentences* in his logic contain "is the true" (the horizontal sign) attached to nominalizations of sentences or other singular terms (like "7"). The horizontal makes any saturated expression into a sentence. Only expressions beginning with the horizontal sign (true sentences) are assertable. The horizontal sign is a component necessary for genuine predication in Frege's logic. In a sense, it is the only genuine predicate in his system. Thus in this subtle way, Frege's logic incorporates the conceptual point that strictly speaking only sentences are assertable. Dummett's famous criticism rests on a straightforward misunderstanding of Frege's theory.

I think that taking "is the true" as a component in all predication—indeed, in a sense, the only sentence-making predicate—is artificial. In fact, I think that there are substantive objections to this view. Frege's deep insight into the structural parallel between singular terms and sentences stands, whether or not one goes along with his view. The structural parallel is that a sentence's truth-value, like a singular term's denotation, is functionally dependent on the denotations of its parts.

I know of no good evidence that Frege ever lost sight of the centrality of sentences—or propositional thoughts—in judgment, inference, language use. Nor did he lose sight of their centrality in his method of reflecting on the semantical structure of language. He is clearly mindful of the deep differences between sentences and singular terms, even as he stresses semantical parallels. Frege's reasons for taking sentences to denote truth-values are complex, pragmatic, and fully in accord with his fundamental philosophical aims. Frege's structural insight is compatible with his insight into the centrality of aiming at truth in language use. It is compatible with the centrality of reflecting on the preservation of truth in inference as the basis for understanding linguistic structure.

Section I of "Frege on Truth" (Chapter 3 below) traces considerations that went into Frege's treatment of truth as the denotation of sentences. It also attempts to assess the relative weight and priority among these considerations. I argue that pragmatic and structural considerations dominate his thinking. This section, which is quite complex, tries to work out details in Frege's method of investigating structure by considering function. He investigates logical and semantical structure by considering language use in relation to the function of judgment. Frege takes the central function of judgment to be that of aiming at truth. He takes the central function of inference to be that of aiming at preservation of truth. Section I is thus an introduction to the background thinking that led to Frege's conceptions of semantics and of logical form.

As I emphasize in section I, Frege's basic considerations in reasoning about the denotation of sentences are structural and pragmatic. His view has ontological implications, however. It is committed to taking truth, the truth-value truth, as an object. This view is widely and correctly regarded as unintuitive, and I do not accept it. Still, I try to show that the view flows fairly naturally from his pragmatic considerations—considerations about how best to account for the cognitively significant aspects of logic and language. Frege's ontological view is also deeply motivated by his belief in logicism. Section III of "Frege on Truth" attempts to elicit considerations that led him to take the truth-value truth as the basic logical object.

With respect to these issues, "Frege on Extensions of Concepts, From 1884 to 1903" (1984) (Chapter 7 below), written earlier, should be read before "Frege on Truth". Relative to these issues, the two papers should be placed in the same Part in this collection of essays. The earlier paper also, however, broaches issues having to do with Frege's epistemology—his rationalism. I placed the paper in Part III of these essays, the section on epistemology, because I wanted to emphasize this aspect of the paper. But the issues of ontology and epistemology are closely intertwined. Here, as elsewhere, the reader would do well to read the essays in different orders and groups, since any order or grouping has its limitations.

The problem of logical objects and the role of truth-values in Frege's account is associated with at least three large, interconnected sources of philosophical interest. One lies in Frege's logicism. Frege's analysis of the structure of arithmetical thoughts and principles led him to believe that the numerals are singular terms. This led him to conclude that they denote objects. This view seems to me natural and probably sound. It raises difficult issues, however, for an account committed to logicism. For it requires not only that logic be committed to objects. It requires that it be committed to what are, to all appearances, particular objects. This commitment seems to be at least prima facie incompatible with the intuition, emphasized by Frege, that logic is completely general—hence not committed to any particular subject matter. It also suggests a problem, raised by Kant: how, if arithmetic is a science whose basic laws are purely general, one can derive conclusions about objects like the numbers, with their particular properties—properties not shared by all objects.[15]

Some writers, both on their own account and as interpreters of Frege, hold that arithmetic is not committed to any particular objects: it is committed at most to structures that any of a number of different objects might fill. This is an interesting position. Discussion of it would take us too far afield. There is some textual basis for the position as an interpretation of Frege. But I believe that the preponderance of evidence counts against it. In my account, especially in section III of "Frege on Truth", I follow the assumption that Frege regarded each number, like the truth-value truth, as a particular object. I think that Frege thought that there is a uniquely natural account of what the logical object truth (the truth-value truth) is. He thought that the particular natural numbers are objects that can be explained in terms of this logical object.

I think that Frege hoped to reconcile this view with the generality of logic. He did so by holding that since truth-values are relevant to and in fact present in any subject matter, since the numbers can be explained in terms of the truth-values as extensions of certain logical concepts, and since logical concepts are applicable in any science, arithmetic is a general science, not a special science. Numbers, like truth-values, are implicit in any subject matter.

Frege's logicism failed. His thinking about what it is to be a general or a special science and his reasoning about what sorts of commitments are universal to any science remain relevant and provocative. This is one of the ways in which Frege's odd-seeming view about truth-values as objects connects to independently interesting issues.

A second source of interest associated with logical objects and the role of truth-values in Frege's account lies in his rejection of the idea that logic is true independently of any relation to a subject matter. This is the idea, mentioned earlier, that the positivists used to exempt logic from epistemic and metaphysical inquiry. Frege saw logic as the most general science of being. This is a traditional view of logic. It stems from Aristotle, saturates the medieval period, and runs through Leibniz, Bolzano, Frege, Russell, and Gödel. I think that it is clearly the dominant view in the history of logic. The view seems to me sound. Explaining in detail, however, exactly how logic relates to a subject matter, and how a subject matter enters into logical truths being true, is a complex and difficult philosophical matter. It seems to me a problem that remains a challenge worthy of serious attention.

[15] Cf. the end of my paper "Frege on Apriority" (Ch. 10 below), for a discussion of this problem. I think that Frege does not fully solve it. His failure to solve it lies partly in the general failure of his logicism. But there are also intuitive, philosophical problems that I believe Frege never fully worked through.

Frege's approach to the problem is radical. He holds that every truth of logic is committed not only to logical functions (which he thought not to be objects) but to the truth-value truth. That is, every truth of logic (indeed every truth) denotes that logical object. Thus truths of logic, like all truths, relate to the world through the simplest and most general semantical relation—the same relation that relates singular terms (or their counterparts in thought contents) to objects. This idea seems to me profound in the way it capitalizes on the ubiquity of truth in framing an account of logic's ontology. Frege's forthright acceptance of the challenge to explain ways in which logic relates to the world also seems to me philosophically profound.

I do not accept Frege's radical approach. I think that the content of ordinary sentences is not correctly understood as a nominalization of the sentence, with a predicate like "is the true" applying to the nominalization.[16] So I do not believe that truth is under discussion, or is denoted, whenever assertions or judgments are made. All assertions and judgments presuppose a commitment to truth. But this commitment is not explicit in simple judgments. In order to make assertions or harbor beliefs, individuals need not have the capacity to make the commitment explicit. I believe that both developmental psychology and epistemology show that attributions of truth represent a cognitive step beyond the mere use of object-level sentences (or thoughts). I believe that these considerations support the view that predication is not strictly functional application, and truth is not strictly the *topic* or *subject matter* of all propositional contents—even the contents of assertions or judgments. Propositional contents have truth-values. But I see no persuasive ground for reifying truth-values, much less taking truth-values to be the topics or subject matter of all propositional representation. Thus I believe that Frege made a mistake in his account of sentences as denoting truth-values. But the mistake does not derive from failure to appreciate his own insight that sentences are deeply different from names.

There are two further grounds for not accepting Frege's radical approach. One is that it leaves one with unacceptably poor resources for understanding the semantical paradoxes, such as the liar paradox.[17] The other is that his redundancy account of the sense of attributions of truth cannot explicate the numerous contexts in which attributions of truth do not apply to fully specified sentences or thoughts ("Berlioz' *bon mot* about Saint-Saëns was true").

[16] Here and elsewhere I use "*thought content*" as the intuitive pre-theoretical notion that corresponds to Frege's theoretical notion of "thought" (in the sense of what is thought). Sometimes I use the term as an alternative to Frege's term, to clarify and reinforce his focus on what is thought, as opposed to the activity of thinking. When I use the term "thought content" on my own behalf, I do not invest it with a commitment to Platonism or to various other doctrines peculiar to Frege. The context will make this clear. I use the term "*representational content*" to apply to thought content, but also to other sorts of content, such as perceptual content, which need not be propositional. Again, I hope that the context will make the usage clear.

[17] For more on the liar paradox, see my "Semantical Paradox", *The Journal of Philosophy*, 76 (1979), 169–198.

Still, Frege's approach seems to me of deep philosophical interest. It deserves much more reflection than standard dismissals have given it.

A third source of interest associated with logical objects and the role of truth-values in Frege's account is the way in which they illustrate Frege's approach to the problem of how we can have apriori knowledge of a subject matter that we do not perceive and bear no causal relations to. I will discuss this issue further in Part III of this Introduction. It is worth indicating here that Frege's account of truth-values as logical objects, discussed in section III of "Frege on Truth", and his pragmatic investigation of the structure of language and logic, discussed in section I of the same paper, are the roots of his approach to these issues.

I mentioned earlier that "Frege on Extensions of Concepts, From 1884 to 1903" (1984) might be seen as belonging in Part I of these essays, as a predecessor of "Frege on Truth"—because of its discussion of Frege's account of structure and ontology. Here is a respect in which "Frege on Truth" should be read as part of the series of papers on Frege's epistemology.

The last paper in Part I of these essays, "Frege and the Hierarchy" (1979), deals with a more special issue in Frege's investigation of structure. It discusses the relation between Frege's theory of indirect denotation and his theory of sense. It attempts to explain principles that Frege was plausibly committed to, and which seem to me plausible in themselves, that generate a hierarchy of senses. He regards increasingly embedded indirect contexts as attributing senses increasingly high in the hierarchy. This early paper of mine is less explicitly historical than the others. I believe, however, that it does interpret Frege's actual position.

A number of significant philosophers, including Carnap, Dummett, and Davidson, have found the hierarchy theoretically unattractive and even untenable. Church developed a version of Frege's hierarchical view. I believe that the hierarchy is not only tenable but almost mandatory. I believe that alternatives for accounting for the relevant linguistic (and conceptual) phenomena have serious liabilities. Understanding these matters is difficult, and may seem to be a rather specialized issue. But I believe that understanding the principles behind the hierarchy not only takes one to the heart of Frege's point of view about sense and thought. The principles seem to me to be fundamental to understanding the functional character and productivity of sense and thought. I also believe that by understanding these principles one comes to have a deeper conception of what sense, as way-of-thinking or mode-of-presentation, is. And this deeper understanding can provide insight into *de re* attitudes toward, and canonical names of, abstract entities. I discuss these issues in a long postscript to the 1979 paper.

The simplicity, depth, and consistency of Frege's method of investigating linguistic and conceptual structure will provide any student of philosophy with valuable insights into language, truth, and thought.

PART II: SENSE AND COGNITIVE VALUE

Frege's notion of sense has undergone waves of critical opposition, from Russell, Wittgenstein, and the positivists, onward. Yet it or close analogs maintain a substantial presence in contemporary theorizing. I want to begin by mentioning two common types of opposition. Neither type seems to me to rest on solid grounds.

One type is rooted in expecting Frege's notion of sense to do explanatory work that it was not intended to do. There is a persistent tendency to see sense as a notion intended to explain ordinary linguistic meaning—the sort of meaning associated with mastery of a communal language. I shall be discussing this source of opposition in some detail in this section. I will not say more about it here.

The other type stems from animus against the ontology of Fregean sense. This animus has a number of sources. I will discuss these briefly.

One source of animus is the assumption that Frege's notions of sense and thought content stand or fall with his Ontological Platonism about them. Recall that Ontological Platonisim is the view that certain abstract entities' existence, or being, and their natures are independent of relations to any entities that exist, or have being, in time. Few philosophers or linguists are comfortable with being as Platonistic about any notion of representational content as Frege is. Although I have come to take Frege's Platonism about the senses or thought contents of some logical and mathematical expressions as a serious and formidable position, I do not accept his Platonism generally—certainly not for representational content about ordinary empirically known entities. I believe, however, that it is a very serious mistake to dismiss his conception of sense because one rejects his Platonic conception of sense. Most of Frege's distinctive theorizing about sense and thought content is valuable independently of the Platonism. What is important is his identification of phenomena that need explanation, the form of explanation that he offers, and the structural and other explanatory insights that derive from his approach.

Many philosophers have maintained that it is a mistake to reify "meanings". This sort of rejection of Frege's notion is often associated with the conflation of sense with linguistic meaning, mentioned above. But it commonly rests on one of three other positions.

One is that there is some irresolvable problem about understanding the relation between speakers or thinkers, on one hand, and the senses or thought contents that they "grasp" or understand, on the other—if such senses or thought contents are "reified". A second is some form of nominalism about representational phenomena (both linguistic and mentalistic). A third is a generalized scepticism toward the objectivity of any discourse about such phenomena.

I cannot undertake to discuss these positions in any detail here. I want to signal my attitude toward them, however. I will remark on them in decreasing order of their ambitiousness.

Generalized scepticism about the objectivity of discourse about meaning, sense, and thought content is associated with Quine, and is sometimes echoed by Davidson. Scepticism is sometimes especially directed at such discourse inasmuch as it discusses sub-propositional "intensional" structure. In my view, this scepticism has never been backed by credible argumentation. It has been ignored, and in my view bypassed, by the empirical development of linguistics and cognitive psychology. No good ground has been given for seeing these sciences as in principle less objective or less genuinely scientific than the natural sciences.

Nominalism about linguistic and representational phenomena was advocated by Ayer, Carnap, Ryle, Austin, Sellars, Goodman, as well as Quine and Davidson. Nominalism about linguistic and representational phenomena is, in my judgment, no better grounded than nominalism about mathematics. Explanation of linguistic and representational phenomena is enhanced by postulating abstractions, as is almost all explanation. To account for shareable linguistic and cognitive phenomena in a simple and uncluttered way, reference to and quantification over representational content that is abstract—not local to any particular place or time—is necessary. Relevant linguistic and cognitive phenomena are the compositionality of propositional understanding, the nature of reasoning and other psychological transformations, shared linguistic and cognitive abilities, the relation between exercises of linguistic or psychological capacities and normative evaluations for veridicality and epistemic success.

I emphasize that not all commitment to abstraction constitutes commitment to Platonism. One might maintain more Aristotelian or conceptualist views about abstract entities. Getting right the exact ontology of the abstractions that are needed to explain linguistic and representational phenomena is complex and philosophically interesting. I believe, however, that it is a secondary matter. The key point about Fregean sense is its explanatory fruitfulness. I believe that one can see, rather easily, that such fruitfulness depends on commitment to abstract entities. But one can make considerable progress by making use of Frege's insights about sense and thought contents without deciding the exact metaphysical nature of the abstractions to which one appeals. This is the analog of the situation with respect to *mathematical* entities in all the sciences. It is fairly obvious that one cannot avoid commitment to such entities or replace them with nominalized counterparts. The metaphysics of mathematical abstractions is, beyond these elementary points, of only tangential importance to explanation that makes use of them. I believe that the ontologies of mathematical abstractions and representational abstractions differ. But in neither case must philosophical and scientific progress await providing settled ontological accounts.

The view that there is an intractable problem about understanding the relation between speakers or thinkers, on one hand, and "reified" senses or thought contents, on the other, can be construed as an instance of a general problem about relations between individual thinkers and abstract objects. I believe that this problem is interesting and important. But I believe that our grounds for believing in the abstract entities are so firm that one can carry out mathematics and the empirical sciences with some confidence without having to solve it.[18]

Sometimes rejection of Frege's position depends on the idea that there is a *special* problem about the relation between individual speakers or thinkers and abstract entities, like representational contents, that the individuals "grasp". The question is how we can get "in touch" with Frege's thought contents (or senses) given that they are abstract entities. This source of ontological animus is more widespread than either of the other sources. It is shared by many philosophers who do not oppose abstract objects *per se*, and who are not sceptics about the objectivity of thought or meaning.

Here I believe that ontological animus against reifying "meanings" or thought contents depends on misconstruing the explanatory role of representational contents. I believe that their role is that of marking shareable states—whether these be linguistic or mental, and whether they be capacities, attitudes, or events. Representational contents play the role of helping to constitute and type-identify explanatory kinds. Representational content is an aspect of a kind of attitude or state. Although the cases are different in several important ways, I believe there is no more difficulty in assuming a representational content associated with a word or a psychological state than there is in assuming a physical parameter or biological kind that has instances type-identified by that parameter or kind. In neither case is there a special problem about explaining the relation between property instances in time and the kind itself.

Sometimes representational contents have been compared to numbers in measurement. This metaphor is misleading in two ways. First, the numbers in measurement can, according to standard representation theorems, be eliminated. By contrast, there is absolutely no solid reason to believe that representational contents can be eliminated from explanations of propositional attitudes (or logic, or semantics). Second, the numbers in measurement are relative to a unit of measure, which is relatively arbitrary. By contrast,

[18] I believe that Frege provides some resources for solving it. I discuss these near the end of this section and in Part III of this Introduction. I also indicate later that I think that the problem arises differently for a mathematical subject matter than for a content of understanding or belief. I think that as applied to contents of understanding and belief, the problem is really a pseudo-problem. When these contents themselves become a subject matter, as in attributions of belief or in semantical theories for thought contents, the problem can seem to re-arise. But I believe that even then, it is different. For I think that canonical names for such contents ultimately depend on understanding of the contents which are the subject matter of these disciplines. So the role of content in understanding—not in reference—is in these cases epistemically basic. For discussion of canonical names, see the Postscript to "Frege and the Hierarchy".

representational contents are fixed by natural, objective relations to a subject matter. They are constituent elements in natural psychological, or other representational, kinds.

Metaphysics can inquire into the relation between attitudes, or individuals and representational contents. But science need not wait for the outcome of such inquiry. Fregean senses play the role of type-identifying linguistic usage, linguistic capacities (including understanding), psychological capacities, psychological states, and psychological events. They also play the role of type-identifying idealized justifications or argument structures. The reason why they should be taken seriously is that they contribute to explanation and to structural insight into the nature of linguistic, psychological, and other phenomena.

Frege himself contributes to the concern that there is some special problem in explaining how we can grasp senses. He sometimes writes as if we have to explain how we can reach over some impossibly large divide:

> . . . the law of gravitation . . . is completely independent of everything which takes place in my brain and of every change or alteration in my ideas. But grasping this law is nevertheless a mental process. . . . in it something comes into account that is no longer mental in the proper sense: the thought; and perhaps this process is the most mysterious of all.[19]

Grasping a thought is simply a misleading metaphor. Any view should cash out the metaphor in terms of having a certain ability to think. Such an ability is attributable on the basis of ordinary evidence and is constitutively associated with a variety of applicational and inferential abilities. How does one "grasp" a thought content? One thinks it. Even on a Platonist view, the thought contents should be regarded as playing a role in type-individuating mental events, states, and abilities. On any view, the contents should be regarded as abstract, in order to account for the multiplicity of instances of events that they type-individuate, to account for the shareability of kinds of thoughts, to account for compositionality, to account for the structure of inference and for various aspects of truth and justification.

Most formulations of the problem of explaining how we grasp an abstract entity, like Frege's formulation, present a pseudo-problem. Often the formulation takes Platonism to regard thought contents as objects that the individual thinker has to target or know about, or as entities that are "before the mind", in something like objects of vision. This amounts to a failure to maintain the distinction between sense and denotation. Such formulations have allowed the grasping or vision metaphors to distort what understanding is. They distort

[19] Frege, "Logic" (1897), in H. Hermes, F. Kambartel, and F. Kaulbach (eds.) *Posthumous Writings*, (Chicago: University of Chicago Press, 1979), 145; cognate passage in the German: *Nachgelassene Schriften*, ed. Hermes *et al*. (Hamburg: Felix Meiner, 1983), 157. Frege goes on, quite plausibly, to ignore the worry, and claim that for the success of scientific enterprises, it is not necessary to solve this "mystery". He notes that it is enough that it be recognized that we do grasp thoughts and recognize them as true.

the explanatory role that thought contents play in type-identifying mental events, states, and abilities.[20]

For the Ontological Platonist, senses or thought contents can and should be regarded as constitutive elements in the individuation of capacities like understanding. The question of how understanding can "make contact with" such entities does not make sense from such a point of view. Understanding is constitutively type-identified in terms of such entities, in terms of thought contents. So the question of how a thinker or understander can "get in touch" with the entities should be, from the point of view of a serious Platonist, ill-formed. The problem of how a thinker can "grasp" the contents reduces to how a thinker can think thoughts with those contents. This latter problem is to be dealt with either through anti-individualist accounts of the individuation of thoughts, or in ordinary scientific ways—through developmental accounts of how an individual learns concepts, or through other psychological accounts of thinking.

As is common in discussions of Platonism, a silly metaphor which is antithetical to the Platonic point of view is substituted for the view. Then its silliness is presented as an objection. Problems with Platonism about understanding senses or thought contents do not arise from its opening a chasm between "realms" which an individual thinker must somehow traverse. Problems arise from reflection on the individuation conditions for kinds of mental capacities. I shall return to these matters.

For now, I want to emphasize that the explanatory structure that Frege develops is more important than his Platonism about representational content. Frege gives his notion of sense four main theoretical functions.[21]

First, sense enters into an account of thought and knowledge. Frege's first introductions of the notion of sense are tied to a notion of cognitive

[20] Michael Dummett makes much of how acute this problem is for Frege given his Platonism about thoughts (*Origins of Analytic Philosophy*, 63–65, 107–108, 135–136). He claims, however, that since Frege takes thoughts to be expressed by language, the problem is not so acute, since it resolves into how one understands a sentence. Much of what Dummett says about this latter problem is plausible. I believe, however, that Dummett is mistaken in thinking that the problem of understanding what it is to think can be solved only if thought is conceived in terms of understanding language. I believe that he is mistaken in assimilating Frege's notion of sense expression to his own conception of linguistic understanding. But primarily I think that he is mistaken in thinking that Frege's Platonism raises a problem about how understanding can traverse a great ontological divide. (Here I distinguish understanding a content from knowledge of a subject matter.) Frege can agree that understanding a thought content just is having certain linguistic or psychological capacities; they are what grasping a thought consists in. Any reasonable view must take the thought contents to be abstract. So all views face ontological questions about the relation between mental events in time and abstract representational kinds and types. But the questions are not those of how a thinker reaches out to thought contents. The ontological issues about abstract thought contents are quite different from issues about how we know mathematical objects. The difference lies in the distinction between expressing a sense and knowing something about a denotation. In the latter case, knowledge takes the objects as denotations or referents. In the case of understanding, the abstract entities are not objects of knowledge. Nor is thought directed to them. Thought is type-identified in terms of them.

[21] For a discussion of functions of sense, see my "Belief *De Re*", *Journal of Philosophy*, 74 (1977), 338–362.

value—thought or knowledge potential. He also characterizes sense in terms of the notion of a mode of presentation—a way of being given to the mind. Senses of declarative sentences are thought contents. Structural parts of sentences express senses that are what are potentially components of thought contents. In theorizing about language, Frege's primary concentration is on a theory of thought and knowledge, as these are expressed through language. The workings of language, the details of linguistic usage, are of interest to him only insofar as they bear on how thought is associated with language in individuals or in communities of individuals. I will elaborate this central point further in this section.

It is important to bear in mind that Frege recognized that there are senses that do not determine a denotation or referent. This is to say that he recognized that thought does not infallibly connect to a subject matter. Names like "Odysseus" or "Ossian" may fail to denote anything. Definite descriptions like "the only round square", may also fail to denote anything. In the normal course of things, senses are ways in which entities determined by the senses are presented to the mind. All senses are representational and thus are essentially associated with a function of determining entities (denotations, objects of reference) presented to the mind. In this sense all senses "purport" to determine a denotation. They are purportedly presentations of a denotation to a mind. But they do not always realize this task or function successfully.[22]

Second, sense has a broadly semantical function, that of fixing or determining a denotation: if a sense determines a denotation, it determines a unique denotation. A sign denotes its denotation and expresses its sense. The sense bears a relation to the denotation of the sign, the relation of *determination*, that is analogous to the sign's relation to its denotation, except that senses, unlike signs, cannot be ambiguous. In an ideal scientific language, Frege held, signs are not ambiguous. They express exactly one sense. Frege thought that in an ideal scientific language signs would never lack a denotation. In such a language, signs would have exactly one denotation, and each sign's sense would successfully determine exactly one denotation.[23]

Frege took it to be a fundamental fact about thoughts—thought contents—that they can be true or false. It is clear that Frege does not regard being true or false as essential to being a thought. Thoughts containing components that are not associated with denotations are neither true nor false. I believe that he thought that successful denotation is the fundamental situation. Thoughts with truth-value have an explanatory priority. But thoughts remain thoughts even when denotation fails and the thoughts are without truth-value.

[22] For a seriously wayward exposition of Frege's view on this last point, see Gareth Evans, *The Varieties of Reference* (Oxford: Clarendon Press, 1982), ch. 1. Evans imposes on Frege a view about object-dependent reference in the senses of proper names that is clearly incompatible with quite a lot that Frege writes—including in "On Sense and Denotation".

[23] Church called the relation of determination the relation *concept of*. Cf. Alonzo Church, "A Formulation of the Logic of Sense and Denotation" in P. Henle, H. M. Kallen, and S. M. Langer (eds.), *Structure, Method, and Meaning* (New York: Liberal Arts Press, 1951).

True and false thoughts are true or false eternally. Thoughts that lack truth-value are also eternally truth-valueless. This is because thought components (including sub-propositional senses) determine an entity, or fail to determine an entity, timelessly. In fact, Frege holds that it is of the essence of thoughts that they are eternal—not in time (or space). Although individual thinkers think thoughts, the thoughts (thought contents) themselves are for him both abstract (not in space or time) and independent of any thinkers.[24] Since senses are essentially potential components of thought contents, they too are in themselves eternal, although they have the role of being senses only in relation to languages, which may well be thinker-dependent.

The claims about eternality are closely associated with Frege's quasi-semantical requirement of unique determination. Senses do not shift denotations with context. Of course, even on Frege's view, what sense is associated with what sign *is* dependent on individuals' mental and linguistic activity. Whether a potential thought component is a sense depends on being expressed by a language. Languages may well be thinker-dependent. So something's counting as a sense may depend on thinkers. But the thought components that are senses (by virtue of being expressed by signs)—the entities that are senses—are themselves, according to Frege, thinker-independent. Thought components and thought contents, which signs relate to when they express them as senses, are not to be explained in any way as dependent on thinkers or language-users. Therein lies Frege's Ontological Platonism. But the semantical function of sense can be separated from this particular metaphysics.

The third theoretical function of sense is to serve as denotation. We can denote the sense of a given sign *S* by using the expression "the sense of *S*". Or we can denote senses in attributions of propositional attitudes. Thus, for example, the sense of a singular term or sentence is, on Frege's view, an object which can itself be denoted and thought about in attributions of propositional attitudes. An attribution like "Al believes that 2 is an even prime" uses a term—the that-clause—that denotes the thought or sense normally expressed by "2 is an even prime"; and the term "2" as it occurs in the that-clause denotes the sense that it would normally express if it were to occur outside any that-clause. Frege provides an elegant and in its main lines plausible theory of how in serious psychological discourse, we can talk about thoughts—both the attitudes and the representational contents.

A fourth theoretical function of sense is that it serve in Frege's account of linguistic understanding. I would like to enter into an extended discussion of this function, since its relation to the other functions—especially the first—is crucial to an accurate understanding of Frege's notion of sense.

Much of Frege's influence on thinking about language derives from his deep structural conceptions. A broader and more general sort of influence

[24] For eternality, see Frege, *Posthumous Writings*, 135. For thinker independence, see ibid. 134. Cognate passages in the German: Frege, *Nachgelassene Schriften,* 146–147.

flowed from his conception of sense, and hence of thought content, as objective and shareable. Frege's emphasis on these aspects of sense crystallized in two famous passages that deeply affected subsequent reflection on language. One passage occurs in section 26 of *The Foundations of Arithmetic* (1884):

> ... we cannot even know whether [space] appears to one man as to another; for we cannot lay the spatial intuition of one next to that of another in order to compare them. But still there is contained in it something objective; all recognize the same geometrical axioms, if only through the deed, and one must do so to find his way around in the world. What is objective in it is the lawful, the conceivable, the judgeable, and what is expressible in words.[25]

The other occurs in Frege's most important and famous essay "On Sense and Denotation" (1892):

> A painter, a horseman, and a zoologist will probably connect very different ideas (*Vorstellungen*) with the name "Bucephalus". The idea is essentially distinguished from the sense of a sign in that the sense can be the common property of many and therefore is not a part or a mode of the individual mind (*Einzelseele*); for one can hardly deny that mankind has a common store of thoughts which are transmitted from one generation to another.[26]

In each of these passages Frege associates what is expressed in words with something shareable among different language-users. Frege's emphasis on the shareability of sense derived from his interest in science. However, he applied it, with qualifications that will emerge, to ordinary use of natural language—as the "Bucephalus" example illustrates.

The positivists and Wittgenstein took over this emphasis. The positivists joined Frege in emphasizing the public character of activity that grounds meaning. They construed the activity that grounds meaning as confirmation procedures in science. The later Wittgenstein took up Frege's remark that recognition of objectivity or shareability in linguistic expression centers "in the deed". Wittgenstein came to criticize Frege's idea that different people might have different uncommunicable intuitions or ideas, a view that Frege developed further in the late essay "The Thought" (1918). But Wittgenstein's work clearly resonated to Frege's view that something objective got expressed "in the deed". Wittgenstein focused on "deed" in actual linguistic behavior and in use in ordinary language.[27] In either application of Frege's

[25] Note that the passage predates the *Sinn–Bedeutung* distinction. I think, nevertheless, that it is relevant to the point at issue.

[26] Substantially the same remark about mankind having a common store of thoughts—for all the multiplicity of languages—occurs in a footnote to Frege's essay "On Concept and Object" (1892).

[27] Contrast Russell's remark: "It would be absolutely fatal if people meant the same things by their words. It would make all intercourse impossible, and language the most hopeless and useless thing imaginable, because the meaning you attach to your words must depend on the nature of the objects you are acquainted with, and since different people are acquainted with different objects, they would not be able to talk to each other unless they attached quite different meanings to their words" ("The

point—to science or to ordinary language—the emphasis on publicity and shareability made investigation of language open to communal and scientific discussion.

The positivists' interest in the publicity of linguistic behavior and the shareability of linguistic meaning was motivated by concern to understand cognitive meaning in scientific enterprises. But the broader application of these ideas to all of natural language use emerged in the work of the later Wittgenstein, Quine, Strawson, and others in mid-twentieth century. This broader application came to be dominant in philosophy. This development in philosophy, along with the Frege-inspired approach to linguistic structure, nourished the emergence in mid-twentieth century of a science of structurally oriented linguistics for "ordinary" or natural language. Thus both Frege's concentration on scientific language and his application of his ideas to natural language contributed deeply to structural thinking about language and to the treatment of language as a public, communal enterprise.

In philosophy the emphasis on public, communal natural language went so far that an important aspect of Frege's thinking came to be obscured. The assumption that we speak a common language (e.g. English) and the assumption that Frege's notion of sense applies only to what is expressed in a common language, to conventional linguistic meaning, led many to the conclusion that Frege had made a number of serious mistakes in his philosophy of language. These mistakes seemed to be magnified as attention turned to aspects of language that he and others thought should be excluded from the language of science. I have in mind context-dependent expressions, vague expressions, simple acts of communication not motivated by scientific concerns. In my view, most such lines of criticism of Frege are based on a systematic misunderstanding of his notion of sense.

I take *conventional linguistic meaning* to be an abstraction that applies to the meaning of expressions in a public language that can be expected to be understood in common by competent speakers of the language. Roughly it is the conventional aspects of the language that are communicated when successful communication occurs. Conventional linguistic meaning is ultimately, I think, an abstraction that is guided by a conception of normal or authoritative competence in a public language.

Frege's conception of sense does overlap modern conceptions of conventional linguistic meaning. Like them, sense is "expressed" by language. It is understood or "grasped" by competent speakers of "the language". It is commonly shareable among different individuals. There is reason to believe that in many cases the sense of a word will be the same as, or very close to the same as, the conventional linguistic meaning of a word. This is especially so

Philosophy of Logical Atomism" (1918), in Bertrand Russell, *Logic and Knowledge: Essays 1901–1950*, ed. Robert Charles Marsh (London: Unwin Hyman, 1956)). Russell's view on this matter—along with his specially restricted epistemology and conception of acquaintance—came to seem quaint, and certainly had little longer-term influence.

when the words are not proper names, demonstratives, or indexicals, and have a relatively non-vague meaning. Even so, the terms in scare quotes ("expressed", "grasped", "the language") must be read with some diffidence. For Frege understands them differently from the way they are commonly understood nowadays.

For all the similarities to modern conceptions of meaning, Frege's conception of sense is deeply different. Frege was firmly centered on understanding the cognitive value of language for a potential science. Thought was more central to him than communication. And both thought and communication were to be understood primarily in the context of reflection on the regimented language of a science. Frege was not specially interested in ordinary use of natural language for its own sake. His genius was such that he had absolutely brilliant insights into the use and structure of ordinary natural language. These insights have led subsequent philosophers and linguists to treat him as an interlocutor, sharing their interests. But reflection on ordinary natural language was for Frege simply a means of gaining insight into the structure of thought, which he wanted a very idealized scientific language to come to mirror and make palpable. Ordinary natural language gave him clues to the structure of thought and thus to the construction of an idealized language.

As a result of Frege's different focus, his conception of the relation between ordinary language use and sense is quite different from modern conceptions of the relation between such use and conventional linguistic meaning. For modern conceptions, ordinary (or at least the most expert) actual use fixes or determines meaning. Meaning is a kind of abstraction from actual usage. For Frege, sense is not fully determined by actual patterns of usage. In addition to linguistic practice (and those aspects of thought that can simply be "read off" such practice), certain ideals of rationality and definiteness played a role for him in delimiting sense. Frege thought that these ideals helped yield insight into the nature of people's actual thinking—a nature that, for him, was not fully fixed by their actual linguistic usage. I believe that Frege was, in the large, right about these points. I shall now proceed to develop them in more detail.

Frege's different conception of sense is signaled most clearly in his discussion of the senses of indexicals in "The Thought" (1918). There Frege completely ignores the obvious point that the conventional linguistic meanings of "I", "today", or a present-tense construction remain constant from occasion to occasion and from speaker to speaker. He maintains that the sense (not just the denotation) of such expressions shifts from context to context. This discussion is clear evidence that Frege's notion of sense does not coincide with any ordinary notion of conventional linguistic meaning.[28] Frege's notion is essentially tied to thought. It is introduced primarily to serve

[28] This conception of meaning is more specific than the very general usage on which sense, denotation, communal linguistic meaning, idiolectic meaning, intended meaning, and so on, are all conceptions of meaning in the broad sense that I discussed in saying, earlier in this Introduction, that

his theory of knowledge. He is interested in language primarily as a means of expressing thought. Communication, communal usage, communal understanding are secondary.

The evidence from Frege's treatment of indexicals is straightforward. The main point of my paper "Sinning Against Frege" (1979) (Chapter 5 below) is to mobilize this evidence to support the contention that Frege's conception of sense differs from modern conceptions of conventional linguistic meaning. Frege's conception is concentrated on thought and motivated by a concern to understand knowledge. "Sinning Against Frege" was for me the beginning of an expanding realization of how markedly Frege's conception of sense differs from common conceptions of linguistic meaning in modern philosophy of language.[29]

Frege's orientation is quite explicit in his first introductions of the notion of sense, in "Function and Concept" (1891).[30] Frege argues first that the equations "$2^4 = 4^2$" and "$4.4 = 4^2$" express different thoughts, though all the signs flanking the equality signs have the same denotation. All denote the number 16. He takes the example to illustrate his view that the thought expressed is to be distinguished from the component denotations. He argues for the same conclusion from the fact that "The Morning Star is a planet with a shorter period of revolution than the Earth" and "The Evening Star is a planet with a shorter period of revolution than the Earth" express different thoughts but have the same denotations. He means by this that an individual can use these sentences as vehicles for thinking different thoughts. He supports his claim that the sentences express different thoughts by noting that someone who does not know that the Morning Star is the Evening Star might accept the one thought and disbelieve or fail to accept the other. Assuming his composition principle that the truth-value of a whole sentence is a function of the denotations of its parts, he supports his claim that the sentences have the same denotation (the same truth-value) by noting that the sentences differ only through exchange of the words "The Morning Star" and "The Evening Star". These names have the same denotation. Frege infers *immediately* from these grounds for distinguishing thoughts and denotations that we should distinguish sense and denotation. In fact, at the end of the paragraph he suggests that the sentences' having different senses just *is* "containing" different thoughts. This train of reasoning makes clear that

Frege made contributions to our understanding of meaning and knowledge. I believe that Frege made contributions to our understanding of all of these conceptions of meaning. But his particular theoretical conceptions, sense and denotation, differ from the others.

[29] Actually, several of the key points in "Sinning Against Frege" already appeared in sect. IV of "Belief *De Re*".

[30] Gottlob Frege, "Function and Concept" (1891), in P. Geach and M. Black (eds.), *Translations from the Philosophical Writings of Gottlob Frege* (Oxford: Blackwell, 1966), 29; "Funktion und Begriff", in Angelelli (ed.), *Kleine Schriften.* (Hildesheim: Georg Olms, 1967), 132. The passage is on page 13 of the original article presented as a lecture to *Jenaische Gesellschaft für Medicin und Naturwissenschaft* (Jena: Hermann Pohle, 1891).

Frege understands his technical notion of sense essentially in terms of its role in representing and accounting for what individuals can think when they use language as a vehicle of thought.

Similarly, in Frege's later but canonical and most famous introduction of the distinction between sense and denotation at the beginning of "On Sense and Denotation", he argues for differences in sense from differences in cognitive value (*Erkenntniswerte*), from differences in what is known, and from differences in epistemic status. Not very much later in the essay, Frege associates sense with mode of presentation (*Art des Gegebenseins*).[31]

Frege does write, three paragraphs further on, "The sense of a proper name is grasped by everyone who is sufficiently familiar with the language . . . ". But one must be careful about what Frege means by "the language". The immediately succeeding passages make it clear that many people may grasp the same sense in a perfect or other scientific language. He indicates that in natural languages, by contrast, an expression may not have a single sense, and that sense may vary with individual speaker and with context. In quite a lot of his work it is clear that Frege takes each person to have his own idiolect, commonly with idiosyncratic senses for proper names and demonstratives used in a context. So contextual ingenuity may be necessary to effect successful communication of certain thoughts.

Frege never clearly states that every sense *must* be graspable by others. He is, however, mainly concerned with shared or shareable senses.[32] He assumes

[31] Gottlob Frege, "On Sense and Denotation", opening paragraph. Cf. *Translations from the Philosophical Writings of Gottlob Frege* 56–57; "Über Sinn und Bedeutung", in *Kleine Schriften* 143–144; the passage is on page 26 of the original article, *Zeitschrift für Philosophie und philosophische Kritik*, 100 (1892).

[32] There is a passage in "The Thought" (1918) where Frege seems to claim that some thoughts are not shareable—certain thoughts expressed through the first-person singular pronoun: in *Collected Papers on Mathematics, Logic, and Philosophy*, ed. B. McGuinness (Oxford: Blackwell, 1984), 359–360; in "Der Gedanke", in *Kleine Schriften*, 350; in the original article, in *Beiträge zur Philosophie des deutschen Idealismus*, 1 (1918/19), 66. This passage is prima facie at odds with passages in which Frege seems to claim that it is of the essence of senses and thoughts (as opposed to ideas (*Vorstellungen*)) that they not only not have bearers but that they are shareable. Cf. e.g. "On Sense and Denotation", in *Collected Papers*, 160; *Translations from the Philosophical Writings*, 59; p. 29 in the original article, *Zeitschrift für Philosophie und philosophische Kritik*; or "Review: Husserl, *Philosophy of Arithmetic*", in *Collected Papers*, 198; in *Kleine Schriften*, 182; pp. 317–318; in the original review in *Zeitschrift für Philosophie und philosophische Kritik*, 103 (1894). The German in these passages is not completely unambiguous, and it is not fully clear that Frege is claiming that the essence of thoughts entails that they be shareable. Frege may be claiming only that many thoughts, or the thoughts at hand, can be shared. It seems to me that charitable interpretation suggests taking Frege not to have contradicted himself or to have changed his mind without flagging the change. It seems to me that Frege's claim about the unshareability of certain particular senses is consistent with the rest of his philosophy. But his Platonism about such senses is extremely implausible for the sorts of reasons that I will discuss below. Dummett discusses the matter in a balanced way in "Indexicality and *Oratio Obliqua*", in *The Interpretation of Frege's Philosophy* (Cambridge, Mass.: Harvard University Press, 1981), 120–4. He regards any claim of unshareability to be part of an "incoherent" doctrine. I disagree with this position. In my view, Frege's remark that there are some unshareable thoughts is not of great importance in understanding his overall view. I believe, however, that it presents a further ground for not identifying his conception of sense with a conception of communal linguistic meaning.

that even in non-ideal non-scientific natural languages, expressions other than proper names, indexicals, and demonstratives have a common sense for most members of a community. So although he sometimes focuses on idiolects, he is also alive to the ways in which individuals share a language, or have significantly overlapping idiolects. It may well be true that for many expressions with a context-free communal meaning, abstracting from vagueness, Fregean sense will be identical with communal meaning.

The relation between communal languages, shared scientific languages, and individual idiolects in Frege's work is a complex one. He never concentrated on the topic. It is often clear, however, that he is concerned with the idiolects of individuals in his discussions of sense. This is particularly true when he is concerned with linguistic expressions that he would regard as not suitable for an ideal language for mathematics—expressions like ordinary proper names and indexicals or demonstratives.[33]

So when Frege states, "The sense of a proper name is grasped by everyone who is sufficiently familiar with the language . . . ", he clearly envisions that "the language" can be either an idiolect, a communal natural language, or an ideal scientific language. If an individual expresses a relatively idiosyncratic sense with a demonstrative or proper name, others may not be sufficiently familiar with "the language" to grasp the contextual or idiolectic sense. In the case of a communal language, the novice or the headstrong might well not grasp a sense shared by most others. In a scientific language those who are not properly trained may not grasp a standard sense. What counts as the relevant language or expression cannot be assumed to be a shared communal language, or an expression with a communally shared sense. Frege considers different types of language, of varying degrees of standardization, depending on the context of his discussion. What matters is what thoughts an individual or group of individuals use signs to think. The sense need not be stable across contexts, even in an individual's language.

I believe that Frege is completely consistent in using his notion of sense as a part of a theory of thought (or thought components) expressed through language. He is completely consistent in taking senses to be modes of presentation, the ways an entity is (actually or purportedly) presented to a mind when one is thinking through language. Language is necessary for the expression of thought. But thought is not, in general, to be explained in terms of ordinary conventional linguistic meaning.

Because of the immense inertia created by the concentration in twentieth-century philosophy on language, especially natural language—more or less in abstraction from its relation to the thoughts of individuals—Frege's notion of

[33] It does not follow that Frege thought that there is some sort of reductive explanation of common languages in terms of idiolects. He does not discuss the matter. There are many possible positions to take on the relation between common languages and idiolects. The only thing that is clear is that Frege thought that some uses of language by individuals express senses that are not communally shared conventional linguistic meanings.

sense was commonly construed simply as communal linguistic meaning, as distinguished from reference or denotation. During the middle part of the twentieth century, in light of the widespread emphasis on the public, communal nature of language, communal usage and understanding in natural language came to be touchstones for assessing the notion of sense. Not surprisingly, Frege's conception of sense does not always happily explain matters that it was not intended to explain. Much of my work on sense has been an exploration of the differences between Frege's conception and notions of meaning that were influenced by his conception but ultimately conflated with it.[34]

The evidence from the passages on indexicals that Frege held a thought-oriented conception of sense that differs from any modern conception of conventional linguistic meaning seems to me decisive.

A rather less straightforward piece of evidence for the same conclusion lies in Frege's unqualified discussion in "On Sense and Denotation" (1892) of different speakers' "attaching" different senses to proper names, even names of individuals well known to the different speakers. Someone interested primarily in communal use and meaning would surely have qualified or explained these remarks. Frege simply takes it as obvious that, at least in natural language usage, different senses will be attached to the same name by different people, as different senses will be attached to the same indexical on different occasions by the same person.

Although Frege gives examples of the senses attached to proper names and indexicals, he never provides a systematic theory. It is likely that he gave no such theory because he thought that individuals could attach any number of senses to a name or indexical. Nothing in the language itself—either as communally spoken language or as a stable idiolect of an individual—constrains the *senses* that might be expressed by such expressions. That is to say that nothing in the language constrains very tightly the thoughts that might be thought through uses of sentences containing such expressions. Frege sees the indexicals as being recruited contextually by thought, as tangible but inexplicit and partial ways of expressing components of thoughts.

In 1970 Kripke and Donnellan presented powerful criticisms of "descriptivist" theories of proper names.[35] In my view, they established a number of

[34] For an example of a fine interpreter struggling with the preconception about sense that I have been discussing, see Dummett, "Indexicality and *Oratio Obliqua*". Dummett writes on the assumption that conventional linguistic meaning or significance is an "ingredient" of Frege's notion of sense (cf. ibid. 100, 125–6, 128, 144). Dummett correctly sees that mode of presentation is an important part of Frege's conception of sense. But he sees Frege as running into tensions or difficulties because these two "ingredients" are in tension. He never provides convincing textual evidence that Frege's notion of sense is centrally concerned with conventional linguistic meaning. So his discussion of problems purportedly internal to Frege's doctrine seem to me to be consistently off base. On the other hand, his discussion of the substantive philosophical issues is subtle and, I think, in important respects right-headed. Cf. also Dummett, *Frege: Philosophy of Language*, 85 ff.

[35] Saul Kripke, *Naming and Necessity* (Cambridge, Mass.: Harvard University Press, 1972); Keith Donnellan, "Reference and Definite Descriptions", *The Philosophical Review*, 75 (1966), 281–304; *idem*, "Proper Names and Identifying Descriptions", *Synthese*, 21 (1970), 335–58.

important facts about the uses of proper names. These facts include the following three: (a) The denotations or referents of most proper names are not fully fixed by descriptions that are associated with the names (either by individuals or in the community). (b) Thoughts associated with the uses of sentences containing proper names are not fully accounted for by citing thoughts that are expressed by definite descriptions substituting for the names. Analogous points apply to indexicals and demonstratives. (c) Causal relations—sometimes running through perception or perceptual memory, sometimes running through the passage of words from one individual to another—play a central role in combining with the individual's usage to fix the referent of names, indexicals, demonstratives, and indeed a variety of other expressions. This last point gained Kripke and Donnellan's initiative the somewhat misleading title "the causal theory of reference". Over three decades later there is still less a theory than a group of very generic principles and powerful reflective observations about particular cases.

Frege's examples of senses of indexicals and proper names all contain descriptions. In every case they also contain further names, indexicals, or demonstratives. Although Frege has been criticized repeatedly as a "descriptivist" about the sense (indeed the conventional linguistic meaning) of proper names, there is no evidence that he maintained any such general theory. That is, there is no evidence that he held a view that if the sense of a proper name (on an occasion of use) were fully spelled out in language, it would in every case be fully expressed in terms of definite descriptions, with no admixture of names, indexicals, or demonstratives. I believe that Frege never thought the matter through in depth.

One way of defending Frege against charges of descriptivism is to appeal to "non-descriptive senses". I think that there probably are such things as non-descriptive senses—non-descriptive intentional thought components or cognitive values that are occasion-of-use-independent. I think that the senses of numerals (contrary to Frege's own view of the senses of numerals) may be examples. But for reasons that I shall discuss in a moment, it seems to me that defenses of Frege along these lines are rather stretched and unconvincing.

There is another consideration that complicates criticisms of Frege from the point of view of the causal theory of reference. "Express" (as in "express a sense" or "express a thought") is a technical term for Frege. There is considerable textual evidence that he regarded the relation *express* between names and indexicals, on one hand, and senses (or thought components), on the other, as very loose. Different people associate different senses with the same name or even with the same indexical in a given set of circumstances. Furthermore, there is no requirement in Frege's theory that when an individual uses a sentence containing a name or indexical on an occasion, the individual thinks only one thought. One might "attach" more than one thought to such a sentence on an occasion of use.

As a substantive matter, it seems to me that this claim can hardly be doubted. So even if (as I believe) proper names almost never express definite descriptions, in the ordinary sense of "express" that characterizes common communal linguistic understanding, it does not follow that they almost never express definite descriptions in Frege's looser conception of "express". For it can hardly be denied that one can think thoughts partly expressible with definite descriptions meant to apply to the denotation of a name, indexical, or demonstrative, when one uses a sentence containing such expressions (names, indexicals, demonstratives).

In view of the looseness of Frege's conception of "expression" when applied to names and indexicals (including demonstratives), one might offer against the standard criticisms the following defense. When one thinks thoughts in using such devices, one often thinks thoughts involving definite descriptions—although these will commonly include further names or indexicals. Who is to say that such thoughts are not among those expressed when one thinks thoughts through sentences containing names or indexicals? One might simply hold that Frege was not specially interested in the sense of "express" in which it has been shown that names (and indexicals) hardly ever express complete definite descriptions.

I believe that any such defense of Frege would be partial, and ultimately insufficient to vindicate his views. One of the upshots of the points made by Kripke and Donnellan is that we think thoughts expressed by sentences containing names, indexicals, and demonstratives, that simply cannot be accounted for by appeal to *complete* definite descriptions. (A complete definite description is a definite description that lacks any context-dependent—demonstrative or indexical—components.) In numerous cases, it is simply not plausible that there are complete definite descriptions available to the individual that could fix the referents of those names, indexicals, or demonstratives.

Thus if one appeals to a "loose" relation of "expression" between these linguistic devices and *complete* definite descriptions, one cannot give a general, psychologically acceptable account of thought. In many cases, there simply is no such definite description in the individual's repertoire.

If one appeals to a "loose" relation of "expression" between these linguistic devices and *incomplete* definite descriptions, the account will be at best partial. For it would not give an account of the fact that often these devices succeed in referring, even though the most salient (incomplete, context-dependent) definite descriptions which the individual associates with the devices do not correctly describe the entity to which the devices succeed in referring. Moreover, in some cases, it appears that no definite description, complete or incomplete, has the same cognitive value—the same potential contribution to knowledge—as the proper name or indexical that it is associated with. These points are fairly well known. So I will not elaborate them here.

Frege's account is subject to a further difficulty in explaining how the senses of incomplete definite descriptions—or any device that relies on context—determine entities (the denotations of linguistic expressions). This difficulty has to do with Frege's commitment to holding that senses determine those entities in a way that is eternal and independent of any thinkers—independent for their individuation of any occasion of use. I shall return to these matters when I discuss their bearing on Frege's theory of thought.

Although I certainly do not accept the full line of defense of Frege that I have just outlined, I believe it to be true that Frege was not interested in conventional linguistic meaning for indexicals and demonstratives. He did not seem interested in stating linguistic rules governing the conventionally understood use of such expressions. This seems completely obvious in his discussion of indexicals and demonstratives. I also believe that Frege was not interested in giving a theory of "meaning", in any ordinary sense of the term, for proper names. So I think criticisms of Frege's philosophy of language regarding indexicals, demonstratives, and proper names are mostly beside the point. His remarks about sense do not line up with most standard purported refutations of his philosophy of language. I believe that errors or oversights in Frege's view, in this area, reside primarily in his theory of thought, not his theory of language. His theory of language is aimed at accounting for different phenomena than modern theories are aimed at.

I will first discuss common types of criticism that rest on misunderstanding. Then I will discuss what I consider to be errors in Frege's views.

Let me go through a common dialectic with respect to Frege's best-known example. Frege claims that "The Morning Star is identical with the Evening Star" requires a distinction between sense and denotation. He states that the component expressions have the same denotations as the expressions in "The Morning Star is identical with the Morning Star", but the two sentences have different senses.[36] Frege reasons that since the senses of "is identical with" and "The Morning Star" remain the same between the two sentences, the senses of the two proper names must be different from one another.

A common objection is to claim that Frege made a mistake in holding that the two names have different senses. Such an objection often holds that names do not have sense or meaning. Sometimes Frege's reasons for his position are not noted at all. Sometimes it is held that Frege made a mistake in *inferring* from differences in knowledge—the difference between knowing that the Morning Star is identical with the Morning Star and knowing that the Morning Star is identical with the Evening Star—to differences in sense. It is held that the differences in knowledge do not reside in differences in the senses of component parts of the sentences.

These lines of objection ignore the fact that Frege understood sense in terms of cognitive value—in terms of contribution to potential knowledge. His notion of sense is introduced to explain differences of knowledge, not some independent set of phenomena. So it is mistaken to represent Frege as

[36] Cf. Frege, "Function and Concept".

making a problematic inference from his point about knowledge to a point about sense. The inference is grounded in his *notion* of sense. This common line of objection does not attend to what Frege's notion of sense is. It assumes without argument a notion of sense that is not Frege's.

It is commonly claimed that names do not have sense because speakers of the communal language do not understand them in the same way. Or it is claimed that Frege made the mistake of taking a name to express the sense of a definite description: any name can be used so as to have distinct sense from any given definite description.

The first of these responses is a *non sequitur*. Frege's notion of sense specifically does not require that names have a common sense in a communal language. Different individuals can express different senses. The senses are normally shareable, but they are commonly not shared. This view derives from Frege's primary concentration on cognitive value, not on communally shared meaning.

Evaluating the second claim is more complicated. In its usual form it depends on an assumption that Frege requires senses to be (complete) definite descriptions. This claim has no basis in the text.[37] Frege does give definite descriptions containing context-dependent devices as examples of senses "expressed" by names. It is incontestable that such definite descriptions express thought components in any ordinary use of a proper name. Given the evident looseness of Frege's notion of sense-expression, it is hard to see these examples as incorporating a straightforward mistake. I think that it would be a mistake to think that the thought content associated with a name is exhausted by such context-dependent descriptions. I believe that names are associated with thought content unlike any of the descriptive contents that Frege cites. Probably Frege failed to realize this. I will come back to this matter. The important point here, however, is that any commitment to holding that senses of names are those of definite descriptions is completely extrinsic to Frege's distinction between sense and denotation for names.

All that Frege's argument depends upon is the requirement that the name be associated on the occasion of use with some thought component (descriptive or not) that is distinct from the denotation of the name. I believe that this requirement cannot be circumvented. In understanding Frege, the important matter is to understand the notion of association or expression that holds between signs and thought components. Frege gives ample ground not to take this relation to be the same as that between signs and conventional linguistic

[37] There are complexities in this case that I will not go into here. It seems to me plausible that any name must be associated by someone (if not the speaker, then someone the speaker depends on) with some descriptive or at least perceptual attributive element, even though the description or attributive element will almost never be complete enough in itself to determine a referent. The interplay between names or demonstrative devices, on one hand, and descriptive elements, on the other, is complex even in communal language use. It is even more complex in thought. For discussion of some of these issues, see my "Five Theses on *De Re* States and Attitudes", forthcoming in a volume in honor of David Kaplan (Oxford University Press).

meanings—shared by competent members of a linguistic community. At least, in the case of names, indexicals, and demonstratives, it is very clear that Frege's remarks that senses are shareable are not to be taken to assimilate senses to communally shared linguistic meanings.

Frege elaborates his distinction between denotation and sense in such a way that a sense can be denoted as well as expressed. That is, one can use language to take sense as a subject matter, not merely as a way of thinking about a subject matter. So the mode of presentation itself becomes the subject matter of a discourse. In some cases, this point seems completely obvious. I can use the expression "the sense of the word 'brown' [or the sense of the numeral '729'] on occasion O". If O is a definite occasion, I can use the whole definite description to denote a certain sense.

Being denoted as well as expressed constitutes the third of the four explanatory functions for the notion of sense that I listed earlier. The explanatory application of this explanatory function can be illustrated as follows. Frege might ask us to consider the attribution, "Al believes that Mark Twain wrote *Huckleberry Finn*". Suppose that this attribution is true. Frege might then note that it is quite possible that "Al believes that Samuel Clemens wrote *Huckleberry Finn*" is false. Al might not know that Mark Twain is Samuel Clemens. Since "Mark Twain" has the same denotation as "Samuel Clemens", exchange of the two expressions in the original sentence ("Al believes that Mark Twain wrote *Huckleberry Finn*") should not change the truth-value of the whole sentence.[38] But there is a common understanding of the sentence according to which the truth-value of the whole sentence does change under this exchange. Or better, the truth-value of a thought commonly thought through use of the new sentence is different from the truth-value of the "old" thought. Frege concludes that the denotation of the two names is not the same *in the context of the sentence and in the way they contribute to the truth-value of similar sentences that contain them.*

Frege maintains (in effect) that in the sentence "Mark Twain is Samuel Clemens" the two names have the same denotation. Note that exchange of the names in this sentence will not change the truth of the whole. But, given his structural analysis that connects denotation with truth conditions, he concludes that *in sentences that attribute propositional attitudes*, the denotations of the two names can differ.[39]

[38] Frege's grounds for believing this principle are discussed in "Frege on Truth" (Ch. 3 below).

[39] It is sometimes held that the name is obviously not ambiguous. It is held that the names "obviously" do not change denotations when they appear in such sentences: "Mark Twain" is still about Mark Twain regardless of what sentence one places the name in. I think that such criticisms are superficial and can be dismissed. Frege can certainly allow for an ordinary sense in which the name continues to be "about" Mark Twain in such contexts. But Frege's notion of denotation, like his notion of sense, is a technical or theoretical one. The notion of denotation is controlled initially by intuitions about reference, but also (and fundamentally) about intuitions about contribution to the truth-value of sentences that contain them. One cannot dismiss Frege's view by appeal to such ordinary intuitions. One must evaluate Frege's theory in relation to its aims and the whole body of evidence that it is meant to explain.

Frege goes on to hold that in the context of the relevant sentences (understood in the way we have understood them), the names denote not their ordinary denotations, but their ordinary senses. Similarly, the that-clauses in the relevant sentences denote not truth or falsity, but their ordinary senses—the thoughts normally expressed by those sentences.

This view results in what I regard as a very beautiful structural account of the way in which the denotation of sentences attributing propositional attitudes depends on the denotations of the parts of those sentences. What I want to focus on here, however, is that Frege takes such sentences to have components that denote the representational thought contents of the thinkers being discussed, including the components of those contents. The sentences are about Mark Twain via being about ways that Al thinks about him. These ways are denoted in the relevant sentences. When language is shared, these ways of thinking are the senses that the expressions within the that-clauses would ordinarily express if they occurred outside any that-clause, for example in a statement of identity.

Here again, a common body of criticism depends on misconstrual of Frege's notion of sense. It is often pointed out that there are readings of sentences like those about Al's belief that do allow for exchange of the names without changing the truth-value of the initial sentence. For example, the reporter might not care how Al thought of Mark Twain. So the report that Al thought that Samuel Clemens wrote *Huckleberry Finn* is just as true as the report that Al thought that Mark Twain wrote *Huckleberry Finn*. In a given reportorial context the latter attribution might not only be true; it need not even be misleading. It is then argued that names lack sense because they are not associated with any commonly understood meaning (much less descriptive meaning). It is further argued that any construal of the sentence that depends on the vagaries of Al's idiosyncratic knowledge base cannot be shared in a communal understanding of the sentence. Sometimes a distinction is drawn between the "semantics" of the name and thoughts pragmatically implicated by the use of the name. It is held that the semantics of a name involves only its referent or denotation. Intuitions that the name might be associated with some other cognitive elements are supposed to be misdirected insofar as they are taken to bear on the name's sense. It is concluded that Frege made a mistake in holding that the names in that-clauses denote the ordinary senses of those names.

In my view, the foregoing line of reasoning is implausible on its own terms. But what I want to focus on here is its relation, or lack of relation, to Frege's own views. Again, Frege's conception of sense is not that of communally understood linguistic meaning. The senses of such names are shareable, but only by understanding the ways in which other individuals think about the names' denotations (or purported denotations). Frege places no requirement on sentences about propositional attitudes in ordinary discourse that the senses denoted be commonly understood or easily identified by others. Frege is also free to recognize the construals of relevant sentences in which senses of names within that-clauses are not denoted, although he does not discuss such sentences and may not have been aware of them.[40] He can account for such construals by holding that they rest on communicative purposes other than those of stating specifically and exactly the nature of an individual's attitudes.

The idea that Frege's notion of sense corresponds to the modern notions of "semantical meaning" or "semantical value", as opposed to pragmatic implicature, derives from the same misconception of the explanatory role of the notion of sense. The idea misses both the orientation of the notion of sense toward explaining cognition and the adjustments one has to make in applying Frege's notion to a non-ideal, natural language. In an ideal language—a language ideally suited to the expression of thought in a science—the sense of an expression would be shared among all competent users engaged in a common scientific enterprise. The sense would also be constant from one occasion to another. Natural language uses of proper names are not like that, according to Frege. In order to capture the sense of a natural language proper name, one has to consider the proper name's use against the central explanatory functions that the notion of sense has. One has to consider the thought associated by the language-user with the linguistic expression on the occasion of use. And one has to consider only cognitive purposes of use—including only the purpose of an attribution in giving the nature of a psychological state, for purposes of psychological explanation, or explanation of the cognitive standpoint of the individual.

Frege's theory applies to explicit and full specification of individuals' thoughts. It is not intended to capture common linguistic understanding. Nor need it apply to all uses of sentences attributing propositional attitudes. I think that there is no question that there is a use of language that is devoted to characterizing peoples' propositional attitudes as such. Finding the appropriate applications of Frege's theory requires distinguishing sense from communally shared linguistic understanding of conventional meaning in a common language. Only in a language devoted to full and explicit characterization of the way others' think—characterization of the specific natures of their propositional attitudes—would Frege's structural scheme for sentences about propositional attitudes have full application. Frege's account is an account of thought expressed and discussed (denoted) in language—not a theory of shared communal usage or understanding.

[40] Thus he is free to recognize that on some occasions, uses of relevant sentences do not attempt to specify fully the individual's thought, and may simply gesture at its general character in such a way as to satisfy practical communicative purposes on the occasion. Language use is quite complex and varied in this respect. On the other hand, I think that there is no plausibility to the view that language use never distinguishes between the truth-values of sentences obtained by exchange of normally co-denoting expressions, or that such use never concerns the ways that individuals to whom attitudes are attributed are thinking. I believe that Frege's strategy is appropriate to *aspects* of many attributions of propositional attitudes in ordinary language. I believe that it is part of a correct theory for language that attempts to talk explicitly and fully about the propositional attitudes of individuals.

Frege regards thoughts, not sentences, as the primary bearers of truth or falsity. In light of this fact, Frege's train of reasoning can be redescribed so as to minimize conflation of sense and conventional linguistic meaning—as follows.

Suppose that we consider attributions of propositional attitudes that are meant to be maximally informative about the nature of the attitude or thought being attributed. These are attributions that in an ideal language would be fitted to scientific, explanatory purposes. In considering non-ideal, natural language, we must center on uses that have primarily explanatory purposes and that are maximally revealing about how the thinker thinks about his subject-matter. Suppose that we assume that we understand and share the thoughts that are thought when another person uses a proper name. In natural language, such thoughts associated with, or "expressed" by, users of a sentence are usually shareable, but they are not always shared. Here we will assume shared understanding. That is, we understand how the person thinks about an object when he or she uses any given proper name; and we think and attribute that way of thinking. Let the sense of an expression be the way of thinking associated with the expression on an occasion of use. The sense of a whole sentence is the thought associated or expressed by the person on an occasion of use.

Mark Twain is identical with Samuel Clemens. The thought component just thought in using "Mark Twain" in the identity sentence just used determines the same denotation—the man Mark Twain—as the thought component just thought in using "Samuel Clemens" in that same sentence. But the thought is conceived as potentially informative, not simply as an instance of a thought of the form "$a = a$".

As a general principle, the truth-value of a thought is functionally dependent on the denotations determined by its parts. So if one exchanges, in a thought, thought components that determine the same denotation, the resulting thought cannot have a different truth-value from the original one.

Now consider an attribution, meant to be maximally informative about Al's psychological state, which is naturally thought when the sentence "Al believes that Mark Twain wrote *Huckleberry Finn*" is used. Suppose that this attributional thought is true. If the denotation that is determined by the thought component (by Al's purported way of thinking) which is associated by us with the occurrence of "Mark Twain" in that attributional use were Mark Twain, that is, Samuel Clemens, then, given the general principle just stated, we would be thereby committed to affirming the thought that is naturally expressed by "Al believes that Samuel Clemens wrote *Huckleberry Finn*". But Al might not know or believe that Mark Twain is Samuel Clemens. In making the original attribution, we are definitely *not* committed to such a thought. So in the maximally informative attributional thoughts, the thought components that we associate with "Mark Twain" and "Samuel Clemens" do not determine the same denotation (the man Mark Twain) that they do in the affirmation of the identity.

We are assuming throughout that the attributional thoughts are meant to be maximally informative about the nature of the attitude or thought being attributed. So it will not do to rest content with an attribution that indicates merely that Al thinks *of* the man Mark Twain that he wrote *Huckleberry Finn*. For no one can think of an individual without thinking of him in a certain way or from a certain attitudinal perspective.[41] We are assuming that we are using language with the purpose of being informative about such perspectives. We want to indicate how Al conceived or thought about that man.

Those thought components associated with the names in the attributions determine ways of thinking that we attribute to Al. The different names are associated with thought components, in the relevant attributions, that determine different ways of thinking. The names in the attributions denote different ways of thinking.

The sorts of thoughts that are thought in uses of these sentences are the sorts that would be expressed in an ideal scientific language devoted to describing and explaining attitudes—specifying those attitudes as fully and as informatively as possible. The notion of "expression" of a sense is fitted to just this sort of relation between language and thought. In this context, it can be postulated that the thought components associated with the that-clauses in the attributional thoughts denote component thoughts attributed to Al. Thought components so attributed are just the thought components that are purportedly expressed by Al when he thinks about an identity. Since senses just *are* cognitive values naturally thought in the use of linguistic expressions, these are the senses of the expressions on those occasions of use. Thus when Al doubts the thought expressed, for him, by "Mark Twain is Samuel Clemens", his thought contains just the thought components that we attribute to him (and denote in our use of those names) when we accept the thought expressed for us by "Al believes that Mark Twain wrote *Huckleberry Finn*" and reject the thought expressed for us by "Al believes that Samuel Clemens wrote *Huckleberry Finn*". Those thought components that are denoted by the different proper names as they occur within the that-clauses are the senses, the thought components, that Al associates with the proper names when he thinks the thought that he associates with the identity sentence. Of course, given that these are ordinary proper names, some idealization is in play. We are assuming a language devoted to a scientific purpose that we share with Al.

The foregoing reasoning is, I think, very close to Frege's. There is more to be said in explanation and defense of this way of looking at matters. But I believe that even as it stands, the reasoning is clearly resistant to the standard sorts of criticism of Frege's account of the structure of attribution of propositional attitudes. In fact, most of those criticisms are clearly irrelevant to this

[41] For discussion of this point, see my "Five Theses on *De Re* States and Attitudes", and "The Content of Perceptual Anti-Individualism", forthcoming.

course of reasoning. Such criticisms rest on presumptions about the "meaning" of sentences that bear no simple relation to Frege's conception of sense, or to its explanatory aims.

I have mentioned that on Frege's view senses are "expressed" by linguistic signs, and are understood by competent speakers of "the language". I have entered significant caveats about the understanding of "expressed" and "the language". Although Frege ignored much that is of interest in the study of the regular use and meaning of natural language, I believe that if one bears firmly in mind exactly what he is trying to explain, most of the standard objections to his use of his notion of sense in his theorizing about language can be seen to be beside the point.

The basic problem with Frege's theory of sense resides in his theory of thought, not his theory of language. Imagine two cases. In one a child forms a perceptual belief. Because of a prismatic distortion he is misperceiving what is in fact a blue ball behind him, and believing of it that it is a red block in front of him. In the second case, a child remembers an object from perception of it. Suppose that she thinks a thought expressed with a demonstrative—"That green ball was fun to play with". This time the child attributes "green ball" correctly to the remembered object. The child does not remember where the object was, or just when she perceived it.

In neither case does the child think meta-thoughts like *the object I am perceiving* or *the object that I perceived*. The child has not yet developed a capacity to think about his or her own perception. Even if he or she were to use such a description, the description need not be capable of picking out the object uniquely on its own. In the first case, the child might be unable to distinguish the time of the perception; the child relies ultimately on demonstratives like *this moment* or *that body*. In the second case, the child might well have perceived other green balls, and has no more than the correct attribution and the memory of the particular ball to single it out.

In each case, it appears completely clear that the child can think about the relevant particular object.[42] But no thought component available to the child that is context-independent (eternal in Frege's sense), and that marks a representational ability that is individuated independently of the particular context of application, suffices to determine the relevant object uniquely.

This is not a type of case that seems to have occurred to Frege. But how is he to account for the child's thoughts? He must postulate a singular sense or thought component, "expressed" by the term "that brown object", which uniquely determines the object. I accept this much. He must also take that

[42] Gareth Evans, *Varieties of Reference*, 90 ff., holds that a memory caused in a normal way by one of two objects that a person once saw, and not caused by the other, cannot refer to the object remembered if the person cannot distinguish the objects in some other way besides memory. I find no plausible defense of this counterintuitive view, despite an elaborate theoretical structure meant to incorporate it. Evans's indefensibly restrictive views on singular reference in thought invade his account of perception as well as memory. I shall discuss these views elsewhere.

sense to be eternal and independent of the thinker, indeed of any thinker. Is there some eternal determining element, some eternal thought component, that applies uniquely to that object, which type-identifies the child's thought?

It is hard to believe in such an element. Neither the children nor the objects are eternal. Each had a beginning and will have an end in time. How could some abstract element of thought determine the object eternally and in a way that is completely independent of the child, and of the perceived or remembered object?

The problem does not lie in the abstractness of the element *per se*. The very idea of a cognitive value or a mode of presentation is one that, as Frege emphasizes, is naturally shareable—at least normally. Shareability is naturally associated with some sort of abstraction from particular instances in individual minds. The problem lies, rather, in the postulated independence of sense, or representational thought content, from thinkers. How is the thought component's intentional or representational character, the fact that it is part of a thought "of" the object, to be explained or explicated?

One can begin to imagine that for each object there are various ways to think about it, fixed by the object's properties. But for any given property there are many ways to think about it, ways that have potentially different cognitive values for a thinker, even one with an ideal understanding of a scientific language. For example, different sense modalities seem to yield different ways of presenting a given property. Different angles of perception on the property within a given sense modality yield different ways. Different capacities within a sense modality (for example, whether a color is represented in a color-blind way or in a normally colored way) can be associated with different cognitive values. A thinker may find informative a thought that constitutes an identification in which one representation of the property is taken to represent the same property as the other. Similarly, for informative thoughts about the identity of objects, or any other type of subject matter. It is hard to see how all of these perspectival modes of presentation could be individuated independently of actual minds and actual objects that individuals interact with.

The problem lies not only in the multiplicity of representational types and their seeming dependence on the perceptual and conceptual capacities of actual individuals or actual species. The problem lies even more fundamentally in the fact that there can be different instances of a given property, or different "duplicate" individuals with indiscernible properties, specified by a given representational type. What instance or individual is perceived, remembered, or thought about, commonly depends on the particular contextual (often causal) relations that the thinker bears to that particular. General representational abilities are often not refined enough to single out the particular by themselves.[43]

[43] These matters are developed in detail in "Five Theses on *De Re* States and Attitudes".

Frege can consistently postulate eternal thinker-independent thought components to correspond to each of these different ways. But such postulation breaks with credibility. Let me summarize separable but closely related difficulties with Frege's theory of thought.

In the first place, the particular content of many thought components makes implausible his apparently generalized commitment to the view that thought components have being independently of thinkers, and indeed of anything in time. Many thought components seem to be contingent on either the nature of thinkers or the nature of other contingently existing temporal entities. It appears obvious, for example, that different degrees of color-sightedness or different types of sense modality are individuated and explained in terms of the mental capacities of thinkers. Modes of presentation in thought can depend on such matters. Similarly, the idea that a conceptual thought component such as that of a harpsichord or a dodo has its being independently of any relation of thinkers to historically contingent individuals of the relevant kinds seems very hard to accept. There seems to be no strong reason to accept it.

In the second place, Frege's view that the representational characteristics of such thought components are, in general, what they are independently of thinkers seems to me to be a piece of magic thinking. It appears obvious that the representational or referential characteristics of such thought components are to be explicated as essentially associated with thinkers' activities or capacities—in terms of actual and potential responses to causal interactions with relevant objects. This point seems to me to apply both to thought components representing repeatable types—such as properties or kinds—and to thought components representing particulars, such as the particular object that the child remembers.

In the third place, Frege's commitment to the view that thought components determine their denotations independently of any particular occasion of thinking—any particular contextual relation between the thought and the denoted entity—fails to be psychologically plausible for numerous cases involving names, indexicals, and demonstratives. The problem is especially acute for names and for demonstratives grounded in perception and perceptual memory. Let me try to make the problem vivid by contrasting the context-dependent thought components with occasion-independent thought components. Let us take as an example of thought content containing only occasion-independent thought components *all beagles are mammals*. The thought component *mammals* represents mammals independently of any particular occasion of thinking with that thought component. Such thought components mark repeatable abilities by a thinker capable of thinking thoughts containing it. The thinker must have causal relations to something (I think usually but not necessarily mammals) in order to represent mammals. But any of various ways of acquiring the concept will do. The ability to represent mammals is repeatable in the sense that it is not essentially tied for its individuation to any particular token interactions with the world.

Frege and many neo-Fregean theories that appeal to "non-descriptive senses" apply his doctrine of occasion-independent determination of represented entities across the board—to every thought component representing every entity. They are committed to maintaining that thought components are fundamentally types, whose identity and representational characteristics are independent of particular events in space or time. So the thought component that determines the object remembered by the child marks a repeatable ability by the child to represent the object—a repeatable ability whose determination of the object is not essentially dependent for its individuation on any particular occasion of thinking about the object. But it is clear that the child's thinking about that particular object can be explained only by reference to the child's causal relations through memory to that particular object. It can be explained only by reference to particular acts of thinking and particular causal relations to that particular object. Generalized repeatable abilities are not sufficient. The child has no way of determining that object independently of particular occasions in which she has been in causal relations (through perception) to it. The representational ability is essentially tied to particular token acts and particular causal relations. Such acts and relations ground modes of presentation that are essentially dependent on particular non-repeatable events in time. I believe that this third point is the most fundamental one.

Some aspects of thought, particularly empirical thought, are irredeemably context-dependent. They are essentially explained in terms of certain particular acts and certain temporally specific causal interactions. Given the limited powers of actual thinkers, some thought about particulars depends irreducibly for its representational identity on certain actual, contextual relations to certain particulars—relations not explainable in terms of general abilities of those thinkers, or context-independent thought components that mark such abilities. I believe that our understanding of thoughts expressed with proper names, indexicals, and demonstratives frequently forces these conclusions. This point seems to me to constitute a sound criticism of Frege. More generally, the explanatory dependence of thought on thinkers' relations to kinds or types of entities—discussed in the first two problems for Frege's view—runs counter to Frege's Platonic position that all thought contents and thought components are thinker-independent.[44]

The problems that I have just discussed are most acute for thought about contingent objects through empirical means. It would be easy for a mathematician to fall into a Platonic view of all thought content. A relaxed Ontological Platonism about mathematical entities is a comfortable and plausible position for a mathematician. It would be easy to transfer Platonism from mathematical entities to mathematical thought contents and then to empirical thought contents, if one is not specifically focused on empirical thought.

[44] For development of these points, often with no special reference to Frege, see my "Belief *De Re*" and "Five Theses on *De Re* States and Attitudes".

Platonism about the contents of empirical thought is much less plausible than Platonism about mathematical or logical objects or functions. Problems of the sort just discussed are also less easy to press for mathematical and logical contents of thought than they are for empirical thought contents. I think that Platonism regarding logical and mathematical thought contents is less implausible. The issues are more complex and puzzling in these areas. I shall not pursue them here.

So far I have distinguished sense from shared conventional linguistic meaning. I have associated sense with a theory of thought through language, not communal linguistic usage. I now want to broach an aspect of Frege's theory of thought that makes his notion of sense even more deeply different from most modern conceptions of meaning.

Frege held that what sense an individual's sign expresses is not fixed by communal linguistic usage and understanding. There is strong textual evidence for thinking that he went much further. It is, in a certain way, not fully fixed either by the individual's own usage and understanding or by use and understanding present anywhere in the individual's linguistic community. This is not a point about the ontology of the entities that are senses. The point is about the relation between linguistic signs and usage, on one hand, and the senses or thought contents expressed by them, on the other. The connection between individuals' expressions (or thought events) and the contents that are thought does depend, at least partly, on the individual's activities and capacities. Individuals express senses and think thoughts only by having certain linguistic and mental abilities, and only by doing certain things. The question is how to understand the relation between these abilities and activities and the expressions of senses or the thinkings of thought contents.

The question can easily invite reduction. Many of Frege's successors attempted to reduce meaning to confirmation procedure, or "use", or functional states defined on the proximal inputs and outputs of individuals. Until the last three decades of the twentieth century, even most who did not propose reductions believed that what "meaning" is expressed by an individual's words is fixed or determined purely by the individual's explicatory abilities and linguistic understanding, fairly narrowly circumscribed.[45]

[45] In the last three decades of the twentieth century, beginning with the work on singular reference by Kripke and Donnellan mentioned earlier, work on natural kind terms by Kripke and Putnam, somewhat later work by me on anti-individualism, this assumption came to be rejected. Cf. Saul Kripke, *Naming and Necessity*; Hilary Putnam, "Is Semantics Possible?" and "The Meaning of 'Meaning' "in *Philosophical Papers* ii (Cambridge: Cambridge University Press, 1975); and my "Individualism and the Mental"; *Midwest Studies in Philosophy*, 4 (1979), 73–121, and "Other Bodies", in A. Woodfield (ed.), *Thought and Object*, (London: Oxford University Press, 1982); repr. in Pessin (ed.), *The Twin Earth Chronicles* (New York: M. E. Sharpe, 1996). The rejection centered primarily on the meaning of expressions in empirical enterprises. Frege's work suggests an anti-individualist view that is mainly centered on considerations in mathematics and logic, although it has, I think, valuable applications to empirical cases. In "The Meaning of 'Meaning' ", Hilary Putnam mistakenly criticized Frege for holding the old view. This criticism seems to me as much dependent on confusions about the new view as on misconstruals of Frege. But it is in line with interpretations of

Frege does not focus on the question. By later standards—associated with attempts to give theories of meaning or explain what meaning "consists in"—Frege has too little to say about understanding and about what determines what sense a word expresses. This is partly a product of his taking sense and denotation as theoretical notions that do not need reduction or philosophical explication so much as fruitful theoretical employment. But there is no question that in light of later concerns, Frege seems relatively uninterested in or oblivious to philosophical issues that cluster around the question. Nevertheless, I believe that Frege had views on the question, and that in some applications, his views are exciting, fruitful, and plausible.

There are a number of places where Frege claims or presupposes that individuals can express senses, or think thoughts through the use of language, that they do not have the background knowledge to fully understand. Such understanding may be lacking even in the individual's broader linguistic community. Frege had particularly in mind cases in mathematics. A relatively vivid example is Newton's thoughts about limits in the calculus. The term "limit" did not receive a fully coherent and adequate mathematical explication until two centuries after its introduction. Frege would have regarded it as having a constant sense and denotation from the time of its initial seventeenth-century uses by Newton. But it came to be fully understood only with the definitions arrived at in the nineteenth century.

The central case for Frege was the term "natural number" itself. Frege thought that no one had adequately understood the term until his logicist explication of it. There is good textual reason for thinking that he thought that the sense of the term "natural number" (or its German counterpart) had remained constant, but that mathematicians had had only an incomplete and infirm understanding of that sense. It is evident that he thought that only with the development of an adequate understanding of "function", and only with the discovery of the functional character of predication, was an adequate understanding of the sense of "natural number" possible.

It is hard to overemphasize how different, in this respect, Frege's conception of sense-expression is from nearly all conceptions of linguistic meaning (or indeed the contents of thoughts) that were developed in the wake of Frege's work. Nearly all other conceptions sought to ground meaning entirely in some sort of actual pattern of use or some sort of actual articulable understanding by the individual, or individuals, whose expressions have the meaning. It is natural to think that either the individual or the community fixes the meaning through an expression's use or understanding.

There is indeed a fruitful conception of meaning that follows these natural thoughts. Frege's conception of sense is, however, different. I believe that it is

Frege (e.g. by Dummett) as mainly concerned with conventional linguistic meaning. For substantive criticism of Putnam—since acknowledged as correct by Putnam—see my "Other Bodies".

also a fruitful one. I believe that it indicates a genuine and important aspect of the nature of thought and linguistic expression.

Earlier, I criticized Frege's conception of sense because of its extreme idealization. I held that its focus on mathematics leads to neglect of particularistic, context-dependent aspects of thought. Here, I believe, we find a respect in which Frege's extremely idealized conception of sense and cognitive value, governed by his extremely idealized conception of an ideal scientific language, yields insights.

Frege believed that norms of reason play a role in determining the nature of an individual's thought and the sense of an individual's (and community's) linguistic expression. He believed that given the function of mathematical thinking and given the fact that mathematical thinking is basically on the right track, the senses of mathematical expressions are partly determined by their role in a correct and uniquely appropriate rational elaboration of actual usage and of incomplete understanding. Thus he regarded his logicist elaboration of ordinary arithmetical discourse as revealing what traditional mathematicians had been thinking with incomplete understanding. It revealed the senses that arithmetical expressions had been expressing all along, even though actual usage and understanding in the past were insufficient in themselves to fix the sense of the expressions. He thought that this elaboration also revealed, of course, the nature of the entities, the numbers, that had been thought and talked about.

In the abstract, this view may seem somewhat far-fetched to those more used to ordinary conceptions of communal or idiolectic linguistic meaning. The failure of Frege's logicism makes this particular application of the view unappealing. But in fact, the view corresponds to something deep about the nature of thoughts. In both the history of natural science and the history of mathematics, there is a strong pull to attributing a conception to individuals who get on to the basic features of a subject matter, even though they have not fully mastered the conception. The example of Newton on *limit* is one such case. The examples of Newton on *mass* or (mathematical) *function*, Leibniz or Descartes on *inertia*, Dalton on *atom*, Mendel on *gene*, Abaelard or Bolzano on *logical consequence*, are others.

Discussion of this aspect of Frege's work is first broached in section VI of "Frege on Extensions of Concepts, From 1884 to 1903" (1984) (Chapter 7 below). It is more fully articulated and developed in "Frege on Sense and Linguistic Meaning" (1990) (Chapter 6 below).[46] The later paper teases out Frege's view from discussion of a puzzle about his views on vagueness. Once one is sensitized to its presence in his writings, it is easy to see it in quite a lot of what he writes. I believe that this discovery about Frege's view is *historically* important, because it provides a much richer background for his conception of sense as fundamentally part of a theory of thought and know-

[46] Cf. also the latter part of my Postscript to "Frege on Truth".

ledge. I regard the account of this aspect of Frege's work as *philosophically* the most important discovery that I have made in working on Frege. I see "Frege on Sense and Linguistic Meaning" as *philosophically* the most significant of my Frege papers.

This view of Frege's is philosophically important, in my judgment, because it makes an important contribution to anti-individualism about the individuation of mental states and about a kind of linguistic meaning-expression. *Anti-individualism* is the view that the nature and the individuation of certain mental states—the correct explication of necessary conditions on what it is to be in certain representational mental states—necessarily involves relations between the individuals in those states and aspects of an environment which is the subject matter of those states.[47] Anti-individualism is an old view in philosophy, going back at least to Aristotle. The view has been more deeply developed in the last quarter-century through its alliance with developments in the theory of linguistic reference (cf. n. 45).

Frege's view is a form of anti-individualism. He holds that an expression can express a definite sense even though the individual's relation to that sense is not entirely explicable in terms of what the individual does with the expression—how he uses it or how he would explicate it. The correct explication of what sense is expressed is supplemented by reference to rationally understandable aspects of reality underlying mathematics and other sciences—aspects that no one, including the individual who expresses the sense, has yet fully understood. This sort of explication is a type of individuation. In other words, Frege holds that sense-expression—what sense a linguistic sign expresses—is individuated partly in terms of the individual's relations to a reality beyond the individual. Frege understands sense in terms of mode of presentation, and sense-expression in terms of the individual's mental states in the use of language. So individuation of sense-expression is individuation of mental states. So Frege's theory of both sense-expression and his theory of mental states are implicitly but clearly anti-individualistic.

My discovery of Frege's application of anti-individualist views to mathematics and his understanding of sense within a rationalist framework made an important contribution to my thinking about meaning, mind, and knowledge. It enriched the development of anti-individualism. It also enriched the development of rationalism, a topic I will discuss shortly.

Frege's rationalist anti-individualism is congenial with his Ontological Platonism about senses—the view that senses themselves are independent for their natures from thinkers. Frege's rationalism is *incompatible* with a traditional *epistemological* Platonism, or at least the traditional caricature of epistemological Platonism. According to this traditional view, understanding is construed in terms of immediate insight, and glossed in terms of the vision metaphor. In mathematics, if what content a thinker thinks or what sense his

[47] Cf. my "Individualism and the Mental" and "Other Bodies".

word expresses is dependent for its individuation on a thinker-independent subject matter of numbers and functions, it is easy and simple to treat the senses as themselves thinker-independent. Then the activity of thinkers plays an individuative role only with respect to what content is thought and with respect to what sense is expressed. It plays no role in individuating the nature of the contents or senses in themselves.

But Frege's views about individuation of sense-expression do not entail Ontological Platonism about thought contents or senses. I believe that these views are best developed in such a way as not to accept a *generalized* Ontological Platonism about senses. Senses are appropriately seen as abstract. But their natures—like the natures of concepts (in the traditional, non-Fregean sense of this term)—are, at least in most cases, better regarded in an Aristotelian or a conceptualist way. It is hard to see a perceptual concept of a beagle, or a concept of Beethoven's Harp Quartet, as independent for its nature of any activity by any thinker. Their abstract identities are not independent of patterns of activity by thinkers in time.

It seems to me that philosophy of language is even now so focused on conventional linguistic meaning that it has not adequately exploited Frege's notion of sense. It has become compartmentalized and scholastic—the negative, flip side of its successful contributions to linguistics. In philosophy the idea that meaning is a common denominator shared among normally competent speakers of a communal language continues to dominate reflection on language. Such a view features the aspects of language and language use that are common to different speakers and common to different occasions of use. Thus the rules governing indexicals, and what communicators have in common (both these rules and the entities referred to on particular occasions) are the center of investigation. Ways in which people think about the entities referred to on particular occasions that are not easily shared or are not incorporated in such rules tend to be neglected. This focus has the advantage of treating language as relatively independent of the psychologies of individuals and vicissitudes of individuals' points of view. The focus has yielded many insights. Frege's role in establishing the focus is fundamental.

I believe, however, that taking better account of Frege's notion of sense will enrich theoretical and philosophical understanding of language. For there is no question that language plays a role in "expressing" the more idiosyncratic aspects of thought that are neglected by the standard modern approach. Frege's notion of sense invites deeper investigation of the complex relation between language and individual thought—and perception—"expressed" in language.

Quite apart from the special issues regarding demonstratives, indexicals, and proper names, Frege's approach highlights certain normative aspects of language use commonly neglected in the focus on conventional linguistic meaning. The historical role of the subject matter and of approximate truth, the role of experimental or theoretical paradigms, and the role of rationality in determining the contents of thought and the senses of linguistic expressions are systematically neglected in the usual approaches to linguistic meaning. These approaches take actual use, or functional relations, or conventional understanding to be determinative of linguistic meaning. But, as we have seen, certain aspects of "meaning" that Frege was responding to in his theoretical notion of sense escape these strictures. Here too Frege's conception promises to supplement our philosophical understanding of language as well as thought.

Linguists in their focus on idiolects rather than communal languages have tended to be more attuned to psychological underpinnings of language than have philosophers. But they have not, in my view, explored linguistic idiosyncrasy or the relation between language and perception very far, much less exploited the anti-individualist and rationalist aspects of Frege's conception. It seems to me that Frege's conception of sense remains a potential source of insight for both philosophers and linguists. I shall return to it in discussing his rationalism.

PART III: RATIONALISM

Frege's conceptions of sense, sense-expression, and sense-understanding must be understood in the context of his larger rationalist views about knowledge. The last group of papers center on Frege's epistemology. Rationalism is the view that some human knowledge has justification or warrant that does not depend for its justifying force on sense experience. Knowledge not warranted by sense experience is usually taken to include logic, pure mathematics, and parts of philosophy. This negative characterization is basic. More positively, rationalism usually maintains that understanding, reflection, or reason suffices to support certain types of knowledge.

Frege was a thoroughgoing rationalist. In some important ways, he was a *traditional* rationalist. He held that knowledge of logic and of the mathematics of number is warranted through reason and understanding. He held that basic principles of these disciplines are self-evident. He thought of them as self-evident not in the sense that they are obvious, but in the sense that they yield in themselves evidence for believing them, if one adequately understands them. Frege was also traditional in believing that there is a natural order of priority among truths in mathematics and logic. This order corresponds to the structure of proof. Frege regarded proof as not purely a matter of establishing what follows from what, but as a matter of justifying a conclusion. This is a traditional view of proof. He thought that the (epistemically) basic truths provide the most fundamental reasons for believing derivative truths. Thus he thought that logic and the mathematics of number rest on a foundation of truths that do not need further justification, because their content provides sufficient "evidence" or justification for belief in them, indeed for knowledge. He held that such foundational truths do not admit of further justification, at least not further ideal or canonical justification. They do not admit justification because their self-evidence is sufficient in itself for knowledge and belief, and not subject to justificational improvement.

With respect to geometry Frege was also a rationalist. In this case he maintained a more complex, purportedly Kantian position. He held that knowledge of Euclidean geometry is warranted through a combination of understanding and pure intuition—a capacity to intuit aspects of space and time without direct use of the senses. Frege applied his traditional views about self-evidence and justificational priority to geometry. His main focus was, of course, knowledge of logic and the mathematics of number.

Frege was aware of the existence of non-Euclidean geometries. He thought of them as mathematical curiosities. He recognized that they are consistent. He maintained that pure intuition shows them to be false.

This position now seems outmoded and difficult to justify. Mathematical practice has validated the non-Euclidean geometries. The notion of pure intuition has remained somewhat obscure. No powerful case has been made that any such capacity shows that non-Euclidean geometries lack mathematical truth. The non-Euclidean geometries gained status through serious arguments that one of them applies to physical space more accurately than Euclidean geometry. Frege could not have foreseen these developments. Still, they tend to undermine his view about non-Euclidean geometries and even his view about the role of pure intuition in supporting any of the geometries. A natural view now, which I believe that Frege would have embraced had he known more, is that at most one of the geometries best applies to physical space; but as pure mathematics, the main geometries are not incompatible. They describe different but genuine geometrical structures. On this view, as pure mathematics, both Euclidean and non-Euclidean geometries are true.

Whether Frege would have held to his view that some type of apriori intuition—some apriori capacity for singular representation—grounds pure geometries is a matter of conjecture. Whether we should regard pure geometry as depending on some such abstract form of singular space-like representation seems to me a question well worth investigation. What is clear is that Frege was mistaken in thinking that Euclidean geometry can be known apriori to be applicable to physical space, and in thinking that it is the only geometry, as pure geometry, that is true.

Frege's error about the non-Euclidean geometries was overshadowed by the spectacular failure of his Logical Law V. Frege proposed this principle as an axiom in order to account for what he regarded as logic's commitment to objects—principally the numbers. The principle maintained a systematic association of extensions with predication. As is well known, Frege's principle was shown by Russell in 1902 to lead directly to contradiction. Russell's demonstration—now known as "Russell's paradox"—defeated Frege's attempt to demonstrate the truth of logicism. Although it contributed to an explosion of attempts either to establish logicism by another route, not least by Russell himself, or to develop a non-logical, set-theoretic foundation for mathematics, it effectively ended Frege's intense commitment to investigating the foundations of arithmetic.

Russell's paradox and the non-Euclidean geometries were long held to have an even more pervasive effect on Frege's philosophy. Logical positivists and others used these examples as bases for claiming that Frege's rationalism, indeed any rationalism, is untenable. The idea was that the two cases showed that rationalist appeals to knowledge warranted independently of experience are completely unreliable.

Not all of Frege's successors accepted such a view. Russell and Gödel did not. But the examples had the effect of giving rationalism a bad name. Euclidean geometry had long been the rationalists' prime example of a case of apriori knowledge—indeed apriori knowledge of physical space. Frege's Logical Law V was often presented by his successors as a principle that seemed perfectly obvious but which turned out not to be true at all. The positivists tended to see the different geometries as competing empirical theories of physical space. As we have seen, they treated logic and pure mathematics as degenerately true—as not about a subject matter at all. Even after the demise of logical positivism in mid-twentieth century, different versions of empiricism hung on. As I have mentioned, it was not until late in the century that empiricism came to be re-examined; and rationalism was taken seriously again.

I believe that the history of geometry reveals a much more complex story than the simple one presented by Frege's successors. The key Parallel Postulate of Euclidean geometry was long regarded by mathematicians and informed rationalists as less obvious than the other principles.

Unlike the Parallel Postulate, Frege's Logical Law V is simply false. It did not undergo or withstand scrutiny for nearly as long as the Parallel Postulate did. The historical story associated with it is again, however, much more complex than the one popularized by Frege's successors.

In "Frege on Extensions of Concepts, From 1884 to 1903" (1984) (Chapter 7 below) I trace the development of Frege's thinking about the key notion, *extension of a concept*, and about Logical Law V. I show that Frege did not take the law to be obvious. He struggled through a complex web of considerations, which even seemed to include attempts to avoid reliance on the "law"—probably because of dissatisfaction with his understanding of the key notion.

This paper is a case study in Frege's particular form of rationalism. What it shows is that despite Frege's claim that basic logical laws are "self-evident", he did not regard all basic logical laws as immediately obvious to a reflective mind. He did not regard self-evidence as subjective or psychological obviousness. He took logical laws to be objectively self-evident and to be subjectively obvious only to a mind that adequately understands them. Like any traditional rationalist, he left room for incomplete understanding that might well lead to mistakes about whether a proposition is a logical law—and even whether it is true.

Most traditional rationalists were fallibilists in this way. They held that purportedly apriori justified beliefs might be mistaken. But traditional rationalist epistemology nevertheless appealed to immediate insight. Many rationalists made claims to knowledge with no more justification than an appeal to self-evidence, in both objective and subjective senses of the term. And often these claims did not withstand the passage of time. Philosophy is, of course, strewn with such claims—later given up, ignored, or seen as mere historical oddities. Frege's empiricist successors often derided the traditional rationalist belief in incomplete understanding as a mere safety valve. They saw it as a kind of dodge that rationalists appealed to, to avoid embarrassment, in case a given claim turned out to have objections to it.[48] The appeal to incomplete understanding was out of step with accounts of meaning in terms of use or confirmation procedures.

The idea that mind is to be partly explained in terms of rational structures that are more fundamental than individual minds is also a traditional rationalist view. This view was largely ignored until very late in the twentieth century. It too was out of step with accounts of meaning in terms of use or confirmation procedures. It took the revival of anti-individualism to make the view seem relevant to contemporary thinking.

In retrospect, the traditional aspects of Frege's rationalism take on different significance in the light of the original elements in his rationalism. There are two broad aspects to Frege's thinking about these matters that are, I think, new in the history of philosophy, and that constitute important contributions. One is, again, his systematic and fruitful connection of issues about mind and knowledge with issues about logic and language. Frege's conception of sense is extremely original in the way it is integrated into a systematic theory of language and thought. His systematic approach to the understanding of thought through understanding language and deductive inference offers the same advantages here as it does elsewhere in his philosophy.

In particular, Frege's holist method for investigating logical form is simultaneously a method for finding and clarifying logical law. Thus full reflective understanding of a logical law is not adequately characterized simply as immediate recognition of the obvious. Any such recognition is the product of a background of competence in inferential practice and structure. Similarly, intuitive understanding of a logical law presupposes an intuitive but discursive competence in carrying out deductive inference. Frege's reflective method clearly depends on the development of theory—not on putative

[48] A skillful representation of this attitude may be found in Imre Lakatos, *Proofs and Refutations* (Cambridge: Cambridge University Press, 1976).

simple, immediate insight. The method presupposes that even unreflective intuitive understanding is ineliminably entangled with a complex web of standing inferential capacities.

Frege's method also allows a wide range of considerations to affect the form and substance of theory—considerations about the functions of arithmetical discourse in counting, for example, or considerations about simplification and unification of different aspects of his logic. Frege's epistemic practice is fundamentally pragmatic. It takes understanding to involve holistic, inferential elements. It is not an epistemology of groundless dogma or of immediate insight.

The other new aspect is his association of his rationalist conception of understanding with the history of science. Major rationalist philosophy before Frege was relatively ahistorical. Frege's historical perspective on the development of logic and mathematics gave him concrete illustrations of incomplete understanding. Frege had the benefit of reflecting on how poorly basic mathematical notions in the calculus had been understood and how they eventually found precise, even definitional, explications. He could rightly regard his development of logic as the clarification of incompletely understood logical concepts and structures. He saw his understanding of number as the culmination of a historical development—attaining a new understanding of a term that had carried its sense through harder, more ignorant times. These cases were not used as dodges. Frege used them to illustrate how incomplete understanding could be clarified and made firm.

Frege said only a little about natural science. What he did say, together with his views on the history of logic and mathematics, makes his conception of sense and his broad emphasis on the possibility of incomplete understanding relevant to understanding the history of empirical science as well as mathematics. Often it seems fruitful and correct to interpret language and thought in natural science in terms of a concept or a meaning that determines a kind before true scientific understanding of the nature of the kind had been achieved. This is true of Dalton's discovery of the atom. Even while he mistakenly defined "atom" as the most basic, indivisible particle, he had a grip on a term that denotes the natural kind. I know of no persuasive reason to think that the sense of his term "atom" was different from ours. Here it appears that the denotation is fixed by relations of the scientific discoverers to the actual kind. The concept or sense expressed is fixed partly through the denotation, partly by approximation to a correct theory or by a fruitful experimental method.[49]

[49] The idea that in the history of empirical science a term should be sometimes understood as expressing a constant concept or sense whose identity is not entirely fixed by current, often mistaken, scientific understanding has been illuminatingly emphasized in various essays by Hilary Putnam, in *Philosophical Papers*, i–ii. The non-historical paper of mine that develops this idea most fully in connection with anti-individualism is "Intellectual Norms and Foundations of Mind", *The Journal*

The two new aspects of Frege's rationalism—logical/linguistic and historical—combined to yield another significant contribution. As I have intimated, he greatly deepened the traditional conception of understanding. Frege's pragmatic approach to understanding linguistic structure, embodying what I called earlier a fundamentally synthetic method, yielded a concrete alternative to the simpler traditional rationalist conceptions of reflective, explicative understanding as simple insight. He went much further than Plato or Descartes in working out a dialectical method that provides and is commonly necessary for "insight". Frege saw such understanding as inextricably associated with systematic, theoretical reflection on a variety of inferences within substantive, scientific theory. So understanding is discursive rather than quasi-perceptual. It is essentially connected to inference. The metaphor of immediate insight is, from his perspective, misleading. What is fundamental for understanding—and ultimately for many rationalist grounds for belief—is connection among inferences. Where understanding provides apriori grounds for belief, it is associated with a range of inferences within an apriori theory, or apriori aspects of theory. Insight is in effect a summation, in understanding, of these articulated inferential capacities.

Frege's historical perspective on the changing understanding of terms in mathematics and empirical science, and especially on the historical development of logic culminating in his own discoveries, led him to construe understanding not as a matter of simply analyzing meanings that were already "implicitly understood". Understanding a sense or a thought is inextricably associated with substantive knowledge. In fact, in some cases, gaining a fuller understanding of one's own thoughts is inextricably associated with *acquiring* new substantive knowledge about other matters. The conditions that individuate concepts or sense-expression lie partly in substantive matters that may not be available to simple reflection, unaided by scientific discovery, no matter how sophisticated the reflection may be. Thus Frege's development of the notion of sense carries him very far not only from the rationalism of intellectual vision. It also carries him very far from empiricist construals of understanding in terms of the unpacking of meaning already implicit in use. It places him even further from the positivist conception of understanding of meaning as "analytically" independent of knowledge of a subject matter.

Frege's philosophical views, and especially his practice as philosopher and mathematician, give fuller substance to rationalist fallibilism and to rationalist appeals to incomplete understanding. I have highlighted Frege's historical perspective and his pragmatic holist method for investigating logical form as key elements in his originality. These elements enabled him to cast rationalism in a new light. Frege's rationalism was quietly embodied in the work of

of Philosophy, 83 (1986), 697–720. This paper is closely associated in its substantive themes with "Frege on Sense and Linguistic Meaning" (Ch. 6 below). The notion of translational meaning in the "Intellectual Norms" paper is close to, and was conceived as, a version of Frege's notion of sense.

Church through much of the century. Rediscovery of its role in shaping his conception of sense and in providing perspective for his conception of incomplete understanding helped inject life into the re-evaluation of rationalism that has been proceeding over the last two or three decades.

"Frege on Knowing the Foundation" (1998) (Chapter 9 below) consists of a detailed examination of both sides of Frege's rationalism—the traditional aspects and the more pragmatic and historical aspects. The paper centers on the meaning of Frege's notion of self-evidence, and on his particular treatment of axioms in his logic.

It seems to me that Frege's rationalism is not vulnerable to criticisms usually put to it by his successors. In the first place, scrutiny of the relevant history indicates that the key examples can reasonably be attributed to incomplete understanding. There was a history of worrying about whether the relevant principles (the Parallel Postulate and Logical Law V) were adequately understood before mathematical developments showed their true status. In the second place, Frege's epistemology does not sanction dogmatic invocation based merely on claims of insight. In the third, there is no good ground for holding that Frege's method of investigation is unreliable or that ordinary non-empirical methods in logic or mathematics are unreliable.[50]

The pragmatic, discursive aspects of Frege's rationalism seem to me to cast Frege's Platonism in an attractive light. "Frege on Knowing the Third Realm" (1992) (Chapter 8 below) discusses Frege's Ontological Platonism and his view of our knowledge of entities that are not in space or time, to which we bear no causal relations, and which are independent for their natures of any thinker.

What I find attractive in Frege's Ontological Platonism is primarily his Platonism about mathematical objects and functions—not his Platonism about senses or thought contents. Frege's Platonism about mathematical entities is for the most part, as I have mentioned, the relaxed Platonism of a working mathematician. There are passages where Frege gives his Ontological Platonism philosophical explanatory work. But for the most part he simply takes mathematics to have an abstract subject matter. He thinks that it is obvious that this subject matter is independent of minds, and other temporal matters, for its nature. As I noted, his extension of his Platonic view to senses and thought components generally is more problematic. Even with regard to his Platonism about mathematical entities, there is an epistemic problem that can seem specially acute for Ontological Platonism.

One of the most famous problems in philosophy is that of explaining how by merely thinking, for example in mathematics or logic, one can know

[50] Of course, philosophy is another matter. I believe that there is apriori knowledge in philosophy. But it is much harder to base such a belief on a history of success than it is to base a belief in apriori knowledge in logic or mathematics. Still, I believe that the situation is less dismal than simple put-downs of philosophy commonly suggest—partly again, because relevant historical accounts are very complex.

something about a subject matter. The problem is usually motivated by noting that in empirical knowledge our thoughts are guided and controlled by causal relations to the subject matter. But in knowledge of abstract subject matters in mathematics or logic, knowledge lacks this guidance and control. So how is it possible merely by thinking (however rationally) to know something about a subject matter? This problem was initially raised by Kant in his question "How is synthetic apriori cognition possible?".[51]

Kant produced an elaborate transcendental idealism to solve his own problem for mathematics and for purportedly apriori cognition in the natural sciences. Kant's theory is philosophically profound and a continuing source of philosophical insight. But, quite apart from its appeal to a putative capacity for pure intuition, its resort to the view that spatial and temporal structures are at bottom mind-dependent seems to me to disqualify it from serious candidacy for being true.

The positivists responded to the problem by trying to protect empiricism. They held that pure mathematics and logic are only vacuously true and are not in any way made true by, and are not true of, a subject matter. They held that all genuine knowledge of a subject matter must be guided and justified by perceptual-causal relations to the subject matter. Kant anticipated this view by holding that logic, though not mathematics, is vacuously true. These appeals to vacuity were, I think, decisively overthrown by Quine in the middle of the twentieth century.

Like the positivists, however, Quine tried to protect empiricism. Against the positivists, he held that truth in logic and mathematics is of a piece with any other truth—true of a subject matter and dependent for truth on a subject matter. He maintained, nevertheless, that logic and mathematics are ultimately known empirically through observations and empirical experiment. They thus remained under the control of causal relations to a subject matter.

This account bears little relation to the actual practice of logic and pure mathematics, and fails to provide a plausible account of the epistemology of these subjects. Pure mathematics and logic rely on understanding and on proof for establishing their truths. They do not look to empirical observations for confirmation. There is no strong reason to think that their practices are not cognitively sound on their own terms. In fact, they are paradigms of cognitively and scientifically powerful enterprises.

This brief catalog of responses suggests, correctly, that the Kantian problem has spawned cures that have been worse than the purported disease. All of these responses in effect give philosophy the role of criticizing, deflating, or limiting the applicability of the sciences of mathematics or logic. One of the sadder themes in the history of philosophy has been philosophy's tendency to

[51] It has been specialized to mathematics and elaborated by Paul Benacerraf in the late twentieth century. Cf. his "Mathematical Truth", *The Journal of Philosophy*, 70 (1973), 661–680. This matter is discussed in some detail in my "Logic and Analyticity".

invoke some less than obvious principle as ground for arrogating to itself the role of dictating acceptable practice, or otherwise radically re-conceiving the main results of sciences whose epistemic credentials are stronger than philosophy's. The epistemic credentials of logic and mathematics are certainly stronger than those of philosophy.

Of course, no recipe limits philosophy's comment on science. Philosophy has sometimes produced critical insights. It has affected the development of science in positive ways. Still, successful global criticisms of scientific practice or successful wholesale re-conceptions of scientific results, especially since Newton, have been rare.

Frege's work suggests an approach to Kant's problem that is refreshingly free of the impulse to criticize or re-conceive mathematical or logical practice. His idea is that a condition on entering into the very practice of these subjects—indeed, on having the very capacity for judgment—is bearing referential relations to a subject matter. Being an individual capable of judging, or having a mind capable of judging, requires "connecting" to a subject matter. It requires getting basic structural aspects of the subject matter right. Frege does not develop his views on this matter. In the discussion of logical objects in "Frege on Truth" (1986) (Chapter 3 below) and in "Frege on Knowing the Third Realm" (1992) (Chapter 8 below) I try to elicit aspects of his thinking that suggest this direction, a direction which I believe to be profound and fruitful.

"Frege on Apriority" (2000) (Chapter 10 below) is motivated by my interest in understanding the history of conceptions of apriority, as well as by a desire to understand certain peculiar aspects of Frege's own conception. The paper traces two conceptions of apriority in Leibniz. One of these conceptions is archaic, but interestingly connected to certain aspects of modern conceptions. The other is the basis for the modern conception. The paper also discusses Kant's refinement of the latter of Leibniz's two conceptions. Kant's conception of experience as sense experience—a conception of experience seemingly more modern (or at any rate more specific) than Leibniz's—makes his conception of apriori warrant or justification substantially the modern one: warrant whose force is independent of sense experience. Although Leibniz and Kant have very similar conceptions of apriority, their background philosophical views lead to very different accounts of how apriority is related to necessity and generality. In "Frege on Apriority", I bring out ways in which Frege's conception is closer to Leibniz's.

Unlike his two great predecessors, Frege is not very interested in necessity. He seems to gloss necessity in terms of apriority. Unlike the conceptions of apriority in Leibniz and Kant, Frege's conception of apriority does not explicitly center on independence of experience for justification. He explicates his notion of apriority *in terms of* a conception of generality: A truth is apriori if "its proof [its basic justification] can be derived exclusively from general laws, which themselves neither need nor admit of proof".[52]

Frege's explication of the notion is defective. It does not explicitly make reference to, or entail, anything about justificational independence of experience. He explicates it in terms of the generality of the principles on which a justification rests. This explication is, I think, vulnerable to counterexamples. What interests me, however, is the way it bears on Frege's views on geometry and arithmetic.

As regards geometry, Frege's apparent agreement with Kant on reasons why geometry is synthetic apriori does not go very deep. Frege's conception of justification in geometry centers on the non-logical character of our capacities to "intuit" space. The justification is non-logical, in that it concerns a special subject matter (space). Kant's conception of justification in geometry centers on the singular character of such intuition. Frege is committed to the justification's being ultimately general.

This difference affects their different views on knowledge (or cognition) of a subject matter. Kant takes synthetic cognition, cognition of a subject matter, to be necessarily grounded in singular thoughts that rest on (singular) intuition. Some intuition is non-empirical or pure, including intuition in geometry. But, for Kant, all substantive theoretical cognition rests for its warrant on singular cognition. Frege takes apriori cognition of a subject matter—hence cognition in logic and mathematics—to rest for its warrant on general cognition. Predication lies at the basis of apriori cognition for Frege. Cognition of abstract objects is warranted through the connection of objects to general predicative representations.

The relative merits of the two views as applied to different sorts of knowledge seems to me to be of great substantive interest. The relation between singular and general elements in representation, cognition, and warrant remains a matter of great importance in contemporary philosophy. It seems to me that both Kantian and Fregean conceptions of warrant and knowledge of subject matters—especially apriori knowledge of abstract subject matters—have significant contributions to make to contemporary thinking.

Frege is a resource for reflecting on numerous central issues in contemporary philosophy. What I have found important and valuable in him differs from what his immediate successors found important, and even from what his best interpreters twenty-five to fifty years ago valued. I believe that Frege's views on semantics, logic, sense-expression, thought, reason, and knowledge are of immediate relevance to current philosophical thinking. I hope that these essays will contribute not only to a better understanding of Frege, but also to creative work—constructive as well as historical—in our great subject.

[52] Gottlob Frege, *The Foundations of Arithmetic*, sec. 3.

1 *Frege (1991)*

Gottlob Frege (1848–1925) lived his life in relative obscurity. He corresponded with some of the great mathematicians of the age, but excepting Russell was largely unappreciated by them. He also corresponded with Wittgenstein, and was heard in lecture by Carnap. These three, particularly Russell, expanded his public and kept his reputation alive. His philosophical importance came to be widely appreciated, however, only in the middle of this century. He is now regarded as the father of analytic philosophy.

Frege founded modern logic. In 1879 he published *Begriffsschrift* which constituted the first fundamental advance in logic since Aristotle. In this work, Frege stated the syntax and semantics for the propositional calculus and first- and second-order quantificational logic. The standard of rigor that he brought to this work was unprecedented.

Frege's work in logic is his greatest contribution. But it was conceived primarily as a means to a further end. He wanted to establish *logicism*, the view that the mathematics of number is reducible to logic. This view derives from Leibniz; but before Frege no one had a sufficiently rigorous or powerful logic to argue for it in a systematic way.

In *Die Grundlagen der Arithmetik*, 1884, Frege set out to define the primary expressions of arithmetic in purely logical terms. The key expressions were '0', 'the successor of', and 'natural number'. Using these definitions he hoped to prove within his logic the theorems of arithmetic. This latter task he reserved to *Die Grundgesetze der Arithmetik* (first volume, 1893; second, 1903).

In *Grundlagen*, one of the most brilliant of all works in philosophy, Frege criticized rival views in the philosophy of mathematics—particularly empiricist, psychologistic, and Kantian views—as a way of motivating the definitions that he proposed. Although his project was conceived as a mathematical one, his genius lay in the deep philosophical motivations that he developed for it. Nearly all of Frege's criticisms of the views he discusses, in the particular forms that he discusses them, are regarded as devastating.

Frege argues that numbers, though abstract and causally inert, are objective. It is disputed whether Frege held the Platonist view that numbers are abstract (not in space and time) and completely independent of minds for their existence and character. But the preponderance of evidence, which grows as his career unfolds, suggests that he was an ontological Platonist, not only about numbers but about functions, thought contents, and various other abstract entities. Frege did not, however, maintain a Platonist epistemology: He did not hold that we have a special intuitive faculty for apprehending abstract objects like numbers. Rather he developed the rudiments of a modern "pragmatic" epistemology, one of his most distinctive philosophical achievements.

The key to Frege's pragmatic epistemology lies in his "context principles", which are stated in various non-equivalent ways in *Grundlagen*. Simplifying, the idea is that one's conception of reference should be derivative from the analysis of the role of expressions (particularly singular expressions) in true propositions. One determines the true propositions, in the usual way, within successful cognitive practice. One identifies successful cognitive practice by seeing what enterprises produce successful communication and reasoned agreement. The content of this doctrine can be seen more clearly in its application to mathematics. Mathematics counts as successful cognitive practice because it yields successful communication, and agreement according to rational, checkable procedures. So its fundamental theorems should be counted true. Given an analysis of the logical form and semantics of truths of mathematics, which Frege carries out, reference to mathematical objects is required for the truth of mathematical theorems. Combined with various arguments that mathematical objects are abstract and mind-independent, the pragmatic epistemology yields a defense of ontological Platonism.

Frege's epistemology rivals a view that would begin by putting constraints on the notion of reference or knowledge (such as the constraint that they have to be accompanied by a causal relation, or be explained in some favored way). Such a view might argue from the claim that mathematical reference or knowledge cannot meet those constraints, to the view that mathematical theorems are not literally true or to the view that mathematics cannot be committed to abstract, mind-independent objects. Frege would regard such a procedure as backwards.

Frege's definitions in *Grundlagen* of the key mathematical terms are very close to those that would be given today. But they rely on the notion of an extension of a concept. In a footnote in section 68 and in section 107 Frege exhibits some unclarity about this notion. Much of his work between 1884 and 1893 was an attempt to clarify the notion, and to justify the key axiom that made use of the notion. This axiom states that all and only F's are G's if and only if the extension of F is identical with the extension of G. This axiom was a key to Frege's logicism. Frege found it less obvious than his other axioms. (Cf. a remark in the Introduction to *Grundgesetze*.)

Again, some of Frege's greatest contributions came as means to a further end. His ground-breaking theory of language in "Function and Concept" (1891), "Concept and Object" (1892), and "On Sense and Denotation"

(1892) was motivated by the desire to clarify the key notion and justify the key axiom. In these articles, Frege proposed to analyze language in such a way that the semantical value of complex expressions would be shown to be a function of the semantical values of their parts. To this end, he took predicates to denote functions. The functions, which he called "first-level concepts", take objects as arguments and yield truth or falsity as values. Higher-level concepts take functions as arguments and again yield truth-values as values. This analysis was the first systematic statement of a compositional, truth-conditional semantics. Such an approach has dominated philosophy of language in this century.

Frege also developed a distinction between sense and denotation (*Sinn und Bedeutung*). The distinction was introduced by an example. In a true sentence of the form "a = b", the denotations or referents of the two proper names are the same. So at the level of denotation, true sentences of that form do not differ from sentences of the form "a = a". But identities of such forms typically differ in what they express; their cognitive values typically differ. Frege proposed that the senses that their respective parts express differ, even though the referents or denotations are the same.

Frege produced parallel compositional theories of sense and denotation. The denotation of a (declarative) sentence was held to be a truth-value. The sense of such a sentence was held to be an abstract thought. The theory of sense enters in an elegant and plausible way into Frege's account of intensional contexts. (An intensional context is a linguistic context in which exchange of expressions that are ordinarily co-denotational does not appear to preserve the denotation of expressions within which the exchange is carried out.) Simplifying slightly, Frege held that in such contexts, expresssions denote their ordinary senses rather than their ordinary denotations, and that substitution of co-denotational expressions *in the context* preserves the denotations of the containing expressions. For example, in "Al believes 2 + 2 = 4", the expression "2 + 2 = 4" denotes not a truth-value, but a thought. Exchange of sentences with the same truth-value as "2 + 2 = 4" will not necessarily preserve the denotation (truth-value) of the whole belief sentence; only exchange of sentences that ordinarily express the same thought will do so—since in the context, "2 + 2 = 4" denotes a thought, not a truth-value. Thus Frege identified some of the primary problems in modern semantics and produced a fruitful and arguably correct strategy for dealing with them.

Frege's notion of sense is less familiar than it may at first seem to be. Although he did associate senses with expressions of natural languages, he did not (or did not in general) identify senses with what moderns would count conventional linguistic meanings. His primary notion for understanding senses was that of a cognitive value, not what is conventionally or normally understood by an expression in a community. He thought that the senses of demonstratives vary with almost each occasion of use, though the conventional linguistic meanings of demonstratives do not thus vary. The idea is that the user's perspective on the world varies with each use. Frege did not believe that sense varies for non-demonstrative expressions to that extent. But he did identify sense with a more idealized conception of cognitive value than would be common today. In fact, he tended to think of senses of non-context-dependent expressions in natural languages as what would be understood by speakers of the language if the speakers had perfected the language for the purposes of knowing about the world (including the world of mathematics) and of expressing that knowledge in an ideally perspicuous way. Thus it was coherent to suppose, from Frege's point of view, that no one could fully and correctly explicate the sense of some expression that is in common use. The sense of an expression might depend on a rationale for its use that is implicit in that use, but that no one has yet come to understand. Thus the sense of number expressions would be fully explicated only when logicism is fully established and articulated. Frege thought that fully understanding (in the sense of being able to explicate) the sense of an expression in a language is not in general separable from understanding the reality that the language is used in knowing.

The sense–denotation distinction remains important in theories of language and cognition. But Frege marshalled it to justify his ill-fated axiom. He developed an intricate argument for claiming that the two sides of the main biconditional in the axiom had the same sense. If this were true, the axiom would clearly be true. But Russell's paradox, which Frege learned of in a letter from Russell in 1902, showed that Frege's axiom is false. This result undermined Frege's version of logicism. Frege's notion of the extension of a concept was never fully clarified. Frege's primary ends were thwarted. But his contributions to logic from 1879 were independent of the axiom. And many of his contributions to philosophy of mathematics and language and to epistemology are of permanent value.

The bibliography of this article is retained as an integral part of the article, unlike the bibliographies in the original publications of the other articles in the collection, whose entries are collected in a bibliography at the end of the collection. The intent is to retain the "introductory" style of this article.

BIBLIOGRAPHY

Burge, Tyler, "Frege on Extensions of Concepts, From 1884 to 1903" *The Philosophical Review*, 93 (1984), 3–34.

—— "Frege on Truth", in L. Haaparanta and J. Hintikka (eds.), *Frege Synthesized* (Dordrecht: D. Reidel, 1986).

—— "Frege on Sense and Linguistic Meaning", in David Bell and Neil Cooper (eds.), *The Analytic Tradition* (Oxford: Blackwell, 1990), 30–60.

Church, Alonzo, *Introduction to Mathematical Logic*, i (Princeton: Princeton University Press, 1956).

Dummett, Michael, "Gottlob Frege", in P. Edwards (ed.), *Encyclopedia of Philosophy* (New York: MacMillan Co. and The Free Press, 1967).

——*Frege: Philosophy of Language* (New York: Harper and Row, 1973).

——*The Interpretation of Frege's Philosophy* (Cambridge, Mass.: Harvard University Press, 1981).

Frege, Gottlob, *The Basic Laws of Arithmetic*, ed. M. Furth (Berkeley: University of California Press, 1964).

——*Collected Papers* (Oxford: Basil Blackwell, 1984).

——*Conceptual Notation* (Oxford: Oxford University Press, 1972).

——*The Foundations of Arithmetic*, trans. J. L. Austin (Oxford: Basil Blackwell, 1980).

——*Philosophical and Mathematical Correspondence* (Chicago: University of Chicago Press, 1980).

——*Posthumous Writings* (Chicago: University of Chicago Press, 1979).

Parsons, Charles, "Frege's Theory of Number", in M. Black (ed.), *Philosophy in America* (Ithaca, NY: Cornell University Press, 1965).

Sluga, Hans, *Gottlob Frege* (London, Boston, and Henley: Routledge and Kegan Paul, 1980).

Part I

Truth, Structure, and Method

This page intentionally left blank

2 *The Concept of Truth in Frege's Program (1984)*

Frege's views on truth are richer and more central to his logicist program than is commonly realized. One reason these views have been underappreciated is that Frege rejected a systematic model-theoretic semantics. Another reason is that many of his views have seemed quaint, naive, or pointless because their underlying motivations have not been pursued in sufficient depth. My purpose in this talk is necessarily quite limited. I want to discuss two cases where too quick a reading of Frege's texts has led to ahistorical assimilation of Frege's views to more modern discussions, to the detriment of an appreciation of his depth as a philosopher.

Frege's argument in "On Sense and Denotation" that a sentence's denotation is its truth-value has often been seen as an elliptical, or even invalid, approximation to an argument to the same conclusion, proposed by Church, Gödel, and others. The Church–Gödel argument, which is clearly inspired by Frege's remarks, presupposes that sentences have a semantical feature that is sufficiently analogous to the central semantical feature of terms to be given the same expression. Let us call this feature "*denotation*". The argument assumes, first, that the denotation of a sentence is not changed by the exchange of co-denotational terms. Second, the argument assumes that logically equivalent expressions have the same denotation. Take any true sentences S and S'; S is logically equivalent with sentence of the form '$(\iota x)\ (x = 0) = (\iota x)\ (x = 0\ \&\ S)$'. So by the second premise, S and this sentence have the same denotation. But the latter sentence yields the sentence '$(\iota x)\ (x = 0) = (\iota x)\ (x = 0\ \&\ S')$' by substitution of co-denotational terms on the right side of the identity sign. So *these* two sentences have the same

The ideas in this presentation are extracted from a paper about ten times as long, 'Frege on Truth' (Chapter 3 below), which is appearing in a volume of the *Synthese Library*. The following works are cited in the body of the paper, and abbreviations used in citing these works by Frege are listed after the title: *Kleine Schriften* (*KS*), ed. I. Angelelli (Hildesheim: Georg Olms, 1967); *The Basic Laws of Arithmetic*, trans. and ed. Montgomery Furth (Berkeley: University of California Press, 1967); *Wissenschaftlicher Briefwechsel* (WB), ed. G. Gabriel, H. Hermes, F. Kambartel, C. Thiel, and A. Veraart (Hamburg: Felix Meiner, 1976); *Die Grundlagen der Arithmetik*, original German included in *The Foundations of Arithmetic*, trans. J. L. Austin (Evanston, Ill.: Northwestern University Press, 1968; Oxford: Basil Blackwell, 1980); *Nachgelassene Schriften* (*NS*), ed. H. Hermes, F. Kambartel, and F. Kaulbach (Hamburg: Felix Meiner, 1968; 2nd edn. 1983). "O" marks the pagination in the original publications of Frege's articles.

denotation, by the first premise (the substitutivity of identity). But the new sentence is logically equivalent with *S′*. So by the second premise, they have the same denotation. So *S* and S′ have the same denotation by the transitivity of identity. An analogous argument shows that all false sentences have the same denotation. Frege accepted not only the conclusion of the argument, but its presupposition and premises. But in arguing for the thesis that the denotation of a sentence is its truth value, he did not advance this argument. Nor did he give an elliptical or invalid approximation to it.

At the stage of motivating his logical theory that his argument is given, Frege treats the "denotation" of a sentence as *whatever is most fruitfully seen as functionally dependent on functional applications among the denotations of sentential parts*. The view that sentences have denotation is thus regarded as a corollary of the centrality of the composition principle in logical theory: the principle that the denotation of a complex expression depends wholly on functional applications among the denotations of semantically relevant parts. Frege's argument that the denotation of a sentence is its truth value is preceded by an argument, which I shall not go into but which I regard as sound, that the sense or thought expressed by a sentence is not to be identified with its denotation.

The argument that a sentence's denotation is its truth value is set out in this passage from "On Sense and Denotation", which I abbreviate slightly:

> The fact that we concern ourselves at all about the denotation of a part of the sentence indicates that we generally recognize and expect a denotation for the sentence itself. The thought loses value for us [at least as scientists] as soon as we recognize that the denotation of one of its parts is lacking . . . But why do we then want every proper name to have not only a sense, but also a denotation? 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 sense to denotation. (*KS* 149; O 33)

Frege's argument has the form: Our interest in the denotation of non-sentential expressions, particularly singular terms, derives from our interest in the truth values of sentences, or in the truth values of thoughts expressed by sentences. Our interest in the truth values of sentences or thoughts derives from our practices of assertion and judgment ("striving after truth"). And it is the business of logic to state the most general laws concerning the norm that governs assertion and judgment. In view of the normative aims of logical theory and in view of the considerations that actually motivate our interest in the denotations of terms, the appropriate feature of sentences to connect with the denotations of non-sentential expressions via the composition principle, is the sentence's truth value. Thus the primary feature of sentences that is of interest to logic, which Frege calls sentence "denotation", is truth value.

Frege's argument is not a deductive argument from "first principles". In "On Sense and Denotation" he twice calls the thesis a conjecture (*KS* 150, 151; O 35, 36). Most of the article is presented as a series of tests of the conjecture—tests of whether truth value can be functionally connected to the denotation of sentential parts. Near the end of the article he suggests that the argument's conclusion has been supported "with sufficient probability" (*KS* 162; O 49).

Unlike the Church–Gödel argument, Frege's does not combine the assumption of substitutivity of identity with the assumption that logically equivalent expressions have the same denotation. Rather it combines a generalization of substitutivity of identity, the composition principle, with the view that logical theory should analyze the basic normative notion governing judgments or assertions and that this normative notion, the notion of truth, motivates our interest in the denotation of sentence parts. Thus the argument unites the point of logic with its basic analytical tool, the composition principle.

Although Frege accepts the second premise of the Church–Gödel argument, the premise that logically equivalent expressions have the same denotation, that premise is less fundamental for him than his views about the point of logical theory and the source of our interest in term denotation. It is virtually apriori true that logical theory ought to count sentences as being the same with respect to their primary logical feature, that is their "denotation", if it counts them logically equivalent. But the notion of logical equivalence that is used in the Church–Gödel argument is explicated in terms of the notion of logical consequence, which is in turn explicated for Church and Gödel in terms of the notion of truth, e.g. necessary truth or truth in all interpretations. Since Frege's argument that the denotation of a sentence is its truth value amounts to an argument that the primary logical feature of sentences is their truth value (and thus that the central notion of logical theory is truth), the second premise of the Church–Gödel argument would be seen by Frege as less fundamental than its conclusion.

Let us turn now to our second example. In section 10 of *The Basic Laws of Arithmetic* Frege identifies the two truth values with specific extensions of concepts. He indicates that the identification is arbitrary relative to the axioms of his logical theory. Any other choice would have been equally consistent with those axioms:

> ... it is always possible to stipulate that an arbitrary course of values [a genus of which extensions of concepts are species] is to be the True, and an arbitrary different one, the False. Accordingly, let us lay down that $\acute{\epsilon}(-\epsilon)$ [the extension of the concept under which only the True falls] is to be the True and that $\acute{\epsilon}(\epsilon = \sim(x)\ (x = x))$ [the extension of the concept under which only the False falls] is to be the False.

This passage has suggested to some authors that Frege's theory of truth had a large stipulative element. Since the numbers are identified with extensions of concepts in ways that presuppose the identifications of the truth values, the remark also suggests that Frege's ontology of the numbers involves substantial and intended arbitrariness or indeterminacy. If these suggestions are correct, then the traditional view that Frege thought that the numbers are genuine abstract objects that can be identified in but one way within his logical theory must be given up. For differently chosen identifications of the truth values with extensions of concepts would produce different accounts of which entities the numbers are. The revised interpretation of Frege based on this passage, which has recently been urged by some authors, would make Frege's position quite analogous to the common modern view that one may identify numbers with sets in any of infinitely many arbitrarily chosen ways.

The passage will not, however, submit to this modernizing interpretation. Frege's identification of the truth-values is indeed arbitrary relative to the axioms of his logic—that is, relative to considerations of consistency. But Frege maintains repeatedly that consistency does not suffice for truth. (Cf. *WB* 75; Frege to Hilbert, 6/1/1900; *KS* 110; O 103; *KS* 264–72; O 321–324, 368–375; *Die Grundlagen der Arithmetik*, 104–119 §§ 92–109). He indicates that philosophical considerations, particularly regarding logicism and his conception of logic have to supplement considerations of mere consistency in order to arrive at a reasonable ontology. (That these points do not appear in *Basic Laws* is consonant with the mathematical orientation of the book. Philosophical points are everywhere, excepting the introduction, severely limited.) Frege's reasoning involves several ideas, but the one most relevant to our purposes is that a semantics for mathematics should not "import anything foreign". That is, it should be in keeping with Frege's philosophical view that mathematics is apriori and reducible to logic. This stricture rules out identifying the truth value with courses of values of empirical or indeed any non-logical (such as geometrical) concepts or functions.

A second restriction emerges from Frege's view of logic as having an internal ordering. Higher-order functions presuppose first-order functions; the functional calculus rests on the propositional calculus. Since the truth values are the "objective" of every sentence of logic and since logic is an ordered elaboration of the laws of truth, the truth values should be specifiable in terms already available in the most fundamental part of logic, the propositional calculus.

Underlying this second restriction is a third. The identification of the truth values as courses of values should be derivative from Frege's concept of truth. Courses of values are derivative, in Frege's view, from functions—and in a deeper sense, from concepts. Frege held a redundancy conception of truth according to which the predication of 'is true' adds nothing to the sense of the sentence to which it is applied. But such predication, though topic neutral, was held by Frege to be ubiquitous among assertions.

Frege identifies truth with the extension of the concept that maps truth onto truth and everything else onto falsity. This concept is represented within Frege's logic by the horizontal. The identification can seem to be artificial or arbitrary if one thinks of extensions of concepts set-theoretically. From this perspective the True is its own unit class. But for a variety of reasons that

I shall not go into, the set-theoretic interpretation misrepresents Frege's standpoint.

Frege's identification of the truth-values derives directly from the preceding meta-logical ideas, together with the view (whose motivations I shall not describe here) that the truth-values are *objects*. The first idea was that the truth-values are *logical* objects. As such their specification must be derivative from specification of logical concepts. The second idea was that logic is an ordered development of the laws of truth, where truth is the aim of sentence use for purposes relevant to logic. Thus Frege takes the logically relevant aim of sentence use to be our "striving after truth", an aim revealed in our practices of assertion and judgment. The object Truth must be involved in the assertive use of every sentence of logic. The two ideas entail that truth is a logical object specifiable in terms of a logical concept that is present in the assertive use of every sentence of logic.

The only concept that is present in the assertive use of every sentence of Frege's logic is the concept denoted by the horizontal—Frege's concept of truth. In unpublished writing contemporaneous with *Basic Laws*, Frege explicitly makes this very point. He writes that what distinguishes 'true' from all other predicates, and what fits it to indicating the aim of logic is that "it is asserted when anything at all is asserted" (*NS* 140, "Logik" (1897)). Assertions or judgments in Frege's system are formulated by adding the vertical judgment stroke to the horizontal.

(The concept denoted by the horizontal meets an analog of a standard condition on truth predicates. Frege specifies this analog in section 5 of *Basic Laws* through the equivalence:

$$\Delta = (-\Delta)$$

where 'Δ' varies over truth-values and may take declarative sentences as substituends. The equivalence is the counterpart in Frege's system of Tarski's truth schema. In specifying the object truth Frege is nominalizing the truth concept by generalizing on his version of the truth schema.)

The critical point is that Frege forces the truth concept to be explicitly denoted whenever a sentence is assertively used in his logic. The aim and subject matter of logic are to be understood only through analysis of assertion and judgment. Frege's identification of the object truth also expresses his conception of the order within logic: The truth concept is the one in terms of which all others are explicated. Identifying truth as a logical object derived from this concept is the only identification that coheres with Frege's philosophical views about the point and epistemology of logic.

The foregoing considerations suggest a rebuttal to two venerable criticisms, made forceful by Michael Dummett, of Frege's claim that truth-values are objects. (Cf. Michael Dummett, *Frege: Philosophy of Language*, 1973, and "Truth" in *Truth and Other Enigmas*, 1978.) One is that once truth becomes one object among others, it is difficult to explain what it is about it that makes us want to strive after it, assert it, acknowledge it, and so forth. The other is that whereas it is part of the concept of truth that we aim at making true statements, Frege's theory of truth as an object leaves this feature of the concept of truth out of account.

I think that these criticisms cannot be sustained. According to Frege, interest in the denotation by terms of objects derives from our interest in norms governing the use of sentences in making assertions and expressing judgments. So using sentences to denote truth-values derives its interest from assertion and judgment. To explain the role of these practices in terms of the features of certain objects, the truth-values, would be to reverse the proper order of explanation.

With regard to the second criticism, it is just not true that Frege "leaves our aim at making true statements out of account" in his articulation of the concept of truth. The specification of the object truth is constructed in such a way as to reflect the primacy of assertion and judgment in revealing the point of logical theory. For all its oddity, Frege's thesis that sentences denote truth-values, which are counted objects, is motivated by a profound conception of the epistemology of logic as rooted in analysis of our practices of assertion and judgment. I believe that understanding Frege's logicist definitions of numerical expressions—hence understanding his logicism—depends on a deeper grasp of the role of philosophical and semantical conceptions in his philosophy of mathematics than has so far been achieved.

3 *Frege on Truth (1986)*

From a natural perspective, Frege's view that sentences denote (*bedeuten*) objects appears to be an irritating peculiarity. His claim that there are only two objects denoted by sentences and that these are Truth and Falsity has seemed to many to advance from the peculiar to the bizarre. Indeed, a standardized form of philosophical humor has grown up around talk of 'naming the True'. I think that the natural perspective is sound and that the humor has its point. But understanding Frege's motivations for these views provides insight into the fundamentals of his philosophical standpoint and method. Such insight enriches the natural perspective.

The importance of Frege's views on truth-values in his system has been appreciated by a number of philosophers. Michael Dummett characterizes Frege's claim that sentences denote objects as 'an almost unmitigated disaster' for Frege's later philosophy of language (*FPL* 196, 643–4).[1] Several

[1] The following works will be cited in the text by the abbreviations mentioned after their titles: Michael Dummett, *Frege: Philosophy of Language* (*FPL*) (London: Duckworth, 1973); Michael Dummett, *Truth and Other Enigmas* (*TOE*) (Cambridge, Mass.: Harvard University Press, 1978); Michael Dummett, *The Interpretation of Frege's Philosophy* (*IFP*) (Cambridge, Mass.: Harvard University Press, 1981); Gottlob Frege, *Begriffsschrift und andere Aufsätze* (*B*), ed. I. Angelelli (Hildesheim: Georg Olms, 1964; 2nd edn. 1977); Gottlob Frege, *Foundations of Arithmetic* (*FA*), trans. J. L. Austin (Evanston, Ill.: Northwestern University Press, 1968; Oxford: Basil Blackwell, 1980); Gottlob Frege, *Kleine Schriften* (*KS*), ed. I. Angelelli (Hildesheim: Georg Olms, 1967); Gottlob Frege, *Translations from the Philosophical Writings of Gottlob Frege* (G & B), trans. and ed. P. Geach and M. Black (Oxford: Basil Blackwell, 1966)—obvious abbreviations of relevant articles from this collection occur in the text; Gottlob Frege, *Die Grundgesetze der Arithmetik* (*GG*) (Hildesheim: Georg Olms, 1962); Gottlob Frege, *The Basic Laws of Arithmetic* (*BL*), trans. and ed. M. Furth (Berkeley: University of California Press, 1967); Gottlob Frege, *On the Foundations of Geometry and Formal Theories of Arithmetic* (*FG*), trans. E. W. Kluge (New Haven: Yale University Press, 1971); Gottlob Frege, *Nachgelassene Schriften* (*NS*), ed. H. Hermes, F. Kambartel, and F. Kaulbach (Hamburg: Felix Meiner, 1968; 2nd edn. 1983); Gottlob Frege, *Posthumous Writings* (*PW*), ed. H. Hermes, F. Kambartel, and F. Kaulbach, trans. P. Long and R. White (Chicago: University of Chicago Press, 1979); Gottlob Frege, *Wissenschaftlicher Briefwechsel* (WB), ed. G. Gabriel, H. Hermes, F. Kambartel, C. Thiel, and A. Veraart (Hamburg: Felix Meiner, 1976); Gottlob Frege, *Philosophical and Mathematical Correspondence* (*PMC*), trans. B. McGuinness and H. Kaal (Chicago: University of Chicago Press, 1980); E. Klemke (ed.), *Essays on Frege* (Kl) (Urbana, Ill.: University of Illinois Press, 1968); Bertrand Russell and A. N. Whitehead, *Principia Mathematica*, i (*PM*) (New York: Cambridge University Press, 1910). "O" marks the original publication, whose pagination is cited, as common coin. Where German editions and translations of Frege's works differ in pagination, both occurrences will be cited, separated by a slash. Responsibility for the translations of all quotations from Frege is mine, although frequently the translations are similar to and benefit from already published translations.

authors have seen in Frege's writings the skeleton of an apriori argument, later given by Church and Gödel, that sentences *must* denote only Truth or Falsity. And Frege's method of identifying the truth-values with certain courses of values has been construed as indicating a non-realistic attitude toward numbers. I think that each of these interpretations is mistaken. But they correctly suggest that Frege's odd-sounding conclusions about truth and falsity should be taken seriously as a key to his philosophies of language, logic, and mathematics.

My aims here are historical. I shall argue in Section I that Frege's view that sentences denote only truth or falsity has profound and natural motivations, and that his view that truth-values are objects is more pragmatically based—and therefore less strange—than has usually been thought. In Section II I criticize Dummett's influential interpretation of Frege's theses on truth-values and his evaluation of the effect of those theses on Frege's philosophy of language. I also delineate the development of Frege's views on assertion and truth between *Begriffsschrift* and *Basic Laws*. In Section III I argue that Frege's identification of the truth-values with the particular objects he identifies them with undergirds his realism about logical objects, and proceeds from some of his deepest philosophical conceptions. In particular, it proceeds from a theory about the nature of logical objects, from a thesis about the aim and ordering of logic, and from his conceptions of assertion and truth.

In order to lay the groundwork for our discussion of Frege's conceptions of assertion, truth, and logical objects, I will have to go over a fair amount of familiar ground in Section I. Some readers may wish to work through this section quickly in order to concentrate on Sections II and III. I should caution, however, that although many of the doctrines discussed in Section I are well known, the ways they fit together and the means Frege uses to motivate them are less well recognized. Understanding these ways and means is critical to a proper appreciation of Frege's use of the notion of truth in his philosophy of logic and mathematics—and indeed, to an appreciation of his depth as a philosopher.

Although defective in various ways, Frege's views on truth are richer and more central to his logical theory and much of his philosophy of mathematics than is often realized. One reason why these views are underappreciated is that Frege refused to allow a meta-theoretic semantics, as we know it, to be an official part of his logical theory. Another reason is that Frege's presentation of his views has tended to encourage concentration on his philosophy of language or his mathematical work as somewhat separate enterprises. The philosophy of language is expounded largely in the great articles of the 1890s and in unpublished writing, with little discussion of its connection to logicism. The mathematical project is spun out in *The Basic Laws of Arithmetic*, which is cast in the form of a traditional mathematical treatise—its philosophy kept to a minimum. Underlying Frege's work is, however, a remarkably integrated vision. We shall try to lay out the central place that Frege's views on truth have in this vision.

I

It is useful to separate Frege's views on truth-values into several theses, although the theses are interrelated and his arguments for them overlap. The relevant theses are

(a) Sentences (when not defective) have denotations (*Bedeutungen*).
(b) The denotation of a sentence is its truth-value.
(c) Sentences are of the same logical type as singular terms.
(d) The denotation of a sentence is an object.

Frege tends to develop support for the theses in the order in which they are listed. (See note 6 below for a qualification.)

Frege's arguments often presuppose his distinction between sense and denotation (which he first draws for singular terms). They almost always presuppose or make use of his groundbreaking composition principles:

(1) The denotation of a complex expression is functionally dependent only on the denotations of its logically relevant component expressions.
(2) The sense of a complex expression is functionally dependent only on the senses of its logically relevant component expressions.

(I omit certain qualifications on these principles that are irrelevant to our concerns.) The first principle is the critical one in Frege's thinking; the second makes important but only occasional appearances.

Thesis (a)

Frege argues that the sense of a sentence—its cognitive value, the thought that it expresses—remains the same regardless of whether or not the sentence's component expressions (particularly, the singular terms) have denotations: The sense of a sentence is fixed independently of its components' denotations [G & B, 'S & R' 63/*KS* 148; O 33; *PMC* 165/*WB* 247; Frege to Russell, 11/13/1904; *PW* 193–4/*NS* 210 'Einleitung in die Logik' (1906).] It follows from (1) and (2) and these considerations that the sense of a sentence cannot be conceived as its denotation.

Frege further argues that one cannot reasonably hold that sentences in general lack a denotation. In at least one passage he draws this conclusion almost directly from the arguments of the preceding paragraph:

It follows that there must be something associated with a sentence that is different from the thought, something for which it is essential whether the parts of the sentence have denotations. This is to be called the denotation of the sentence. (*PW* 194/*NS* 210–11; 'Einleitung in die Logik' (1906))

This inference clearly relies on the Composition Principle (1).

Now one might feel that the inference begs any question one might have about whether sentences have denotations—about Thesis (a). Why should sentences be included among the complex expressions that have denotations? The last sentence of the passage just cited suggests something wrong with the question. Frege is not using the term 'denotation' with a fixed, self-understood meaning in arguing for (a). Rather he is determined to give his Composition Principle (1) a comprehensive role in logical theory, and he is intending to fit the term 'denotation' to the role that the principle might fruitfully play in a logical theory about sentences. 'The denotation' of a sentence is whatever is most fruitfully seen as functionally dependent on the denotations of its parts. So far the phrase 'the denotation' has no specific logical grammar or ontological implications. Since the arguments for (a) do not presuppose Thesis (d), there is so far no reason to consider the view that a sentence's denotation is an *object*. One may at this point regard talk about sentence denotation as potentially a *façon de parler* for an important semantical aspect of sentences. The ontological import of such talk, if any, is left thoroughly open.

Of course, the term 'denotation' ('*Bedeutung*') was not devoid of intuitive content in Frege's arguments. '*Bedeutung*' is a common word in German, usually translated 'meaning'. In German there is no oddity in saying that sentences have a '*Bedeutung*'. Frege did, however, appropriate the term for his theoretical uses and introduced it in the essays 'Function and Concept' and 'On Sense and Denotation' through examples of singular terms ('The Evening Star', 'Odysseus'—which lacks a denotation—'1', '2 + 2', 'the capital of England'). The examples suggest that naming or reference—considered as relations between names and their bearers or between a complex singular term and the object it picks out—is one primary sort of '*Bedeutung*'. But since Frege also used his term to apply to a semantical relation between expressions (such as predicates), that he emphatically did not regard as singular terms, on one hand, and non-linguistic entities, on the other, one must view these initial examples with some caution. They are aids in building a theory.

The point I want to press regarding Frege's quick inference to (a) from his composition principle is that the inference is indicative of his pragmatic attitude toward his terminology. As he repeatedly noted, the term 'number' had expanded in its application (from the natural numbers, to negatives, rationals, reals, complex numbers) under pressure from the requirements of mathematics. Semantical terminology could be expected to undergo similar stretching in response to the demands of logical theory.

Frege provides a closely related, but different argument for (a). This argument occurs in 'On Sense and Denotation' and is repeated in his correspondence with Russell and his posthumously published writings. The following passages suggest the argument:

> The fact that we concern ourselves at all about the denotation of a part of the sentence indicates that we generally recognize and expect a denotation for the sentence itself. (G & B, 'S & R' 63/*KS* 149; O 33)

> Now it would be impossible to see why it was of value to us whether or not a word had a denotation if the whole sentence did not have a denotation and if this denotation was of no value to us; for whether or not that is so [whether or not the words have a denotation] does not affect the thought. (*PMC* 152/*WB* 235; Frege to Russell, 12/28/1902)

> [If a sentence had no denotation] the denotation of any part would be a matter of indifference, for, regarding the sense of a sentence, only the sense not the denotation of its parts comes into consideration. (*PMC 158/WB* 240; Frege to Russell, 5/24/1903; cf. also *PMC* 165, 163n/*WB* 247, 245n; Frege to Russell, 11/13/1904, *PW* 232/*NS* 250–1; 'Logik in der Mathematik' (1914))

These claims are embedded in discussion of examples of nondenoting names and in an argument for Thesis (b). But in view of the obvious generality of their intent, I think that they are worth isolating.

Frege's argument is that we would not concern ourselves with the denotations of sentence-parts if we were not interested in the denotations of whole sentences; we clearly do concern ourselves with the denotations of sentence-parts—we often care whether singular terms denote something; so we are interested in the denotations of whole sentences.

The argument must again be seen in the light of the centrality of the Composition Principle (1) and of Frege's pragmatic use of the term '*Bedeutung*'. Denotations of sentences are whatever can be seen as both central to logical theory and functionally dependent on the denotations of the logically relevant parts of sentences. But this argument adds a further claim. Our interest in the denotations of words is derivative from our interest in the denotations of sentences. That is, word denotation is important because and only because of the importance of some feature of sentences that is central to logical theory and functionally dependent on word denotation.

This further claim appears to be an expression, or outgrowth, of the context principles that Frege had enunciated earlier in *The Foundations of Arithmetic*. These enunciations preceded the development of the distinction between sense and denotation, and they took a variety of nonequivalent forms. But they all emphasized that the analysis of the "meaning" of a word (in retrospect, presumably, its sense and its denotation) was to be carried out in the context of an analysis of its role in a sentence. Frege appears to be invoking the primacy of sentences in his argument that sentences have denotations. Our interest in the denotations of words had to be connected in logical and linguistic theory with some feature of sentences. Frege forged the connection by means of the Composition Principle (1), and he called the relevant feature of a sentence its denotation.

It is worth noting that Frege's reasoning is prima facie *incompatible* with the idea that the notion of the denotation of a term has no other content than that provided by an analysis of the contribution of the term in fixing the denotation (or truth-value) of a sentence.

The argument presupposes that we have a co-equal understanding of and application for the notion of the denotation of a term.[2] Indeed it presupposes that the notion of term denotation is more familiar than that of sentence denotation, though perhaps not more familiar than that of truth-value. The argument claims that whether terms have any denotation at all is of importance to us only relative to our interest in relevant semantical properties of sentences. It does not suggest that the notion of term denotation can be exhaustively defined, or characterized, or reduced by attempting to analyze the relevant semantical properties of sentences in total abstraction from one's ordinary understanding of the notion of term denotation (reference). The ordinary understanding of term denotation is assumed to be sound. (One could produce numerous passages from Frege's opposition to formalism to substantiate this point.) The argument simply demands that such ordinary understanding has to be connected, in one's theory, to the semantical properties of sentences, interest in which motivates interest in the denotations of terms.

Thesis (b)

The role of value and "interest for us" in Frege's argument for (a) needs articulation. Frege saw logic as revealing certain norms governing ideal thought. The sentence was the linguistic correlate of thought. We think, according to Frege, only by means of sentences. So any logical theory had to ground itself in an analysis of the properties of sentences that revealed the relevant norms. Our interest in the denotations of terms, and of functional expressions, was motivated by interest in normative properties governing thinking, normative properties whose laws logic sought to uncover.

[2] The fact that Frege's notion of term-denotation cannot be entirely separated from the 'name-bearer' relation has been appropriately emphasized by Michael Dummett in *FPL*, chapter 12; and in *IFP*, chapter 7. For some views that proceed on the assumption of the primacy of sentences (or their truth-values) and on the view that the notion of the denotation of a term has no content other than that which is derivative from an analysis of how the term functionally determines truth-value, see W. V. Quine, *Word and Object* (Cambridge, Mass.: MIT Press, 1960), chs. 1–2; *idem*, *Ontological Relativity* (New York: Columbia University Press, 1969), chs. 1–2; J. Wallace, 'Only in the Context of a Sentence do Words have Meaning', *Midwest Studies, 2 (1977);* Donald Davidson, 'Reality without Reference', *Dialectica*, 31 (1977), 246–58; Hilary Putnam, *Reason, Truth, and History* (Cambridge: Cambridge University Press, 1981), chs. 1–2. Several of these authors explicitly invoke Frege's inspiration. I find the view not only uncongenial to Frege (though unquestionably inspired by part of his doctrine) but unpersuasive. But I shall not be able to go into these points here.

This focus on the normative implications of logical theory underlies Frege's primary argument for Thesis (b). Frege specifies what it is about sentences that motivates our interest in the denotations of their parts. The relevant property is the sentence's truth-value.

The fact that we concern ourselves at all about the denotation of a part of the sentence indicates that we generally recognize and expect a denotation for the sentence itself. The thought loses value for us as soon as we recognize that the denotation of one of its parts is lacking. We are therefore justified in not being satisfied with the sense of a sentence, and in asking also for its denotation. But why do we then want every proper name to have not only a sense, but also a denotation? Why is the thought not enough for us? Because, and to the extent that, we are concerned with its truth-value. This is not always the case. In hearing an epic poem . . . we are interested only in the sense of the sentences and the images and feelings thereby aroused. In response to the question of truth we would abandon aesthetic delight and turn to a scientific investigation. Hence also it is a matter of no concern to us whether the name 'Odysseus', for example, has denotation so long as we accept the poem as a work of art. It is the striving for truth that drives us always to advance from sense to denotation. (G & B, 'S & R' 63/*KS* 149; O 33)

When we merely want to enjoy the poetry we do not care whether, for example, the name 'Odysseus' has a denotation . . . the question first acquires an interest for us when we take a scientific attitude—the moment we ask, 'Is the story true?', that is, when we take an interest in the truth-value. . . . Now it would be impossible to see why it was of value to us whether or not a word had a denotation if the whole sentence did not have a denotation and if this denotation was of no value to us; for whether or not that is so does not affect the thought. And this denotation will be something that will have value for us precisely when we are interested in whether the words have denotation (*bedeutungsvoll sind*), therefore when we ask after truth. (*PMC* 152/*WB* 235; Frege to Russell, 12/28/1902)

. . . if it is not a matter of indifference to us whether the signs that make up a sentence have a denotation, then it is not just the thought that matters to us, but also the denotation of the sentence. And this is the case when and only when we ask after truth. Then and only then does the denotation of the sentence enter into our consideration; it must therefore be most intimately bound up with truth. (*PMC* 165/*WB* 247; Frege to Russell, 11/13/1904)

That the name . . . designates is of value to us when and only when we are concerned with truth in the scientific sense. So our sentence will have a denotation when and only when the thought expressed in it is true or false. (*PW* 232/*NS* 250–1; 'Logik in der Mathematik' (1914))

Frege's argument for Thesis (b) clearly presupposes his arguments for Thesis (a). It thus presupposes the primary importance of sentences in logical theory. In fact, the language of the first three passages directly echoes the first statement of context principle in *The Foundations of Arithmetic*: 'never to ask for the *Bedeutung* of a word in isolation, but only in sentential context'. The phrases I have translated 'asking for its denotation' and 'ask after truth' in these three passages from 'On Sense and Denotation' and the correspondence with Russell use the same phrase '*zu fragen nach der Bedeutung*' that occurs in the introduction to *Foundations*. It is almost inconceivable that Frege did not intend to associate the passages with his slogan. The *reason* that the denotations of words must be 'asked for' only in sentential context is that the relevantly related semantical feature of sentences—the denotation of sentences—motivates our interest in word denotation. Our interest in the denotation of words derives from our interest in the truth-value of sentences, or of the thoughts that they express. Truth is the relevant norm governing our use of and interest in sentences and thoughts. The point of logical theory should be the analysis of the most general laws governing this norm.

Frege's argument for Thesis (b), the thesis that the denotation of a sentence is its truth-value, is not and is not intended as a deductive argument. There is no attempt to deduce (b) from 'first principles'. In 'On Sense and Denotation', he twice calls the thesis a conjecture (*Vermutung*—conjecture, supposition, surmise)—(G & B, 'S & R' 64, 65/*KS* 150, 151; O 35, 36). And the remainder of the article is presented as a series of "tests" of the conjecture. After considering these tests in detail, he writes at almost the end of the article: 'From this it follows with sufficient probability that the cases where a subordinate clause is not replaceable by another with the same truth-value proves nothing against our view that a truth-value is the denotation of a sentence whose sense is a thought' (G & B, 'S & R' 78/*KS* 162; O 49).

There is a closely related argument, proposed by Church, Gödel and others, that does take deductive form. Frege has sometimes been constructed as giving an elliptical, or even invalid, approximation to this argument. I think that such a construal is very poor history. Frege's argument rather has the form: In view of the normative aims of logical theory and in view of the considerations that actually motivate our interest in the denotations of terms, the appropriate feature of sentences to connect with the denotations of the sentence's constituent parts via the Composition Principle, is the sentence's truth-value. We shall discuss the Church–Gödel argument shortly.

There is a supplementary argument for Thesis (b). This argument is roughly: a sentence's truth-value is dependent on the denotations of its constituent parts in just the way that the Composition Principle (1) requires. So given the way the notion of a sentence's denotation is introduced, truth-values are well suited to be the denotations of sentences.

In 'On Sense and Denotation' and elsewhere Frege proposes this argument as a confirmatory consideration or an essential test of the conclusion of the previous argument: 'If our conjecture that the denotation of a sentence is its truth-value is correct, the latter must remain unchanged when a part of the sentence is replaced by an expression having the same denotation. And this is in fact the case' (G & B, 'S & R' 64/*KS* 150; O 35).

Taking the appeal to the Composition Principle (1) as a confirmation or supplement to the previous argument seems to me to be Frege's most reasonable presentation of the relation between the two arguments for Thesis (b). But sometimes Frege seems to place the appeal to the Composition Principle in a different light. He asks, "what else but the truth-value could be found, that belongs quite generally to every sentence, to which the denotation of its constituent parts is relevant, and that remains unchanged by substitutions of the kind in question?" (G & B, 'S & R' 64/*KS* 150; O 35; cf. also *PMC* 158, 165/*WB* 240, Frege to Russell, 5/24/1903; 247, Frege to Russell, 11/13/1904).

Although Frege cannot be expected to have foreseen this, his question prompted Russell to open a semantical and metaphysical Pandora's box. One can well imagine Russell turning over in his mind this question, which Frege put to him more than once in their correspondence of 1902–4. For Russell was resisting the view that sentences have truth-values as denotations.

A year after the correspondence ended, Russell published his theory of descriptions. The theory opened the possibility of maintaining allegiance to the Composition Principle (1), yet analyzing the logically relevant parts of a sentence in a very different way from the way Frege regarded as natural and appropriate. The theory simultaneously opened the possibility of assigning a variety of different sorts of denotations to sentential parts, and a variety of different sorts of denotations, other than truth-values, to the sentences themselves ('states of affairs', 'facts', 'propositions' and so forth).

Russell demonstrated that one could do compositional semantics without taking truth-values to be the central feature functionally associated with sentences. But it is no accident that, despite the deep methodological interest of the theory of descriptions, Frege's approach, not Russell's, has been the source of the mainstream development of semantical theory in logic. No doubt one reason for the pre-eminence of Frege's approach lies in the artificiality, from a syntactical or grammatical point of view, of Russell's analysis of sentential constituents. But more profound reasons are suggested by Frege's first argument for Thesis (b). Truth (or some modalized notion of truth, like necessary truth or validity) is the central notion of logical theory. In making truth-values the primary functional values of the Composition Principle (1), Frege was simply uniting his formal apparatus with the conception that motivates logical theory.

Russell's theory can, of course, accommodate the representation of truth-values, of truth-evaluations; and it maintains allegiance to a notion of logical consequence explained in terms of truth. But formally, the truth-values enter through a side door, so to speak. The composition principle yokes words with sentences, but it is not used primarily to relate word-denotation to truth-value. The primary semantical feature of sentences is the 'fact' they are correlated with. The denotations of words functionally determine a 'fact' or 'proposition' composed of attributes and (perhaps) individuals. The primary semantical feature of a sentence is the 'fact' that it is correlated with. Thus, from the outset, Russell's formal theory incorporates into its subject matter entities that evince a strong admixture of metaphysical motivation. States of affairs, facts, and the like have a recurring attraction for the metaphysically minded. But they have not obtained general acceptance among logicians, and they have yet to be shown to be indispensable for the foundations of logic. By contrast, the more abstract notion of truth is firmly entrenched in nearly all logical theories. Formal logical theories that place this latter notion at their center, resting little or no weight on arguably dispensable metaphysical entities semantically correlated with sentences, have formed the mainline development of logic in this century. Frege may be seen as a certain sort of minimalist in this context. He conceived of the fundamental part of logic—the calculus of truth-values and first- and second-order logic—as having an aim and subject matter that is relatively independent of metaphysical controversy. The laws of logic are fundamentally the laws of truth, not laws about the metaphysical constitution of facts, propositions, or thoughts (*Gedanken*). (Cf. Kl 508/*KS* 342; O 58; and *PW* 122/*NS* 133 'Ausführungen über Sinn und Bedeutung' (1892–1895).) The metaphysics of thoughts is developed to deal with intensional contexts and with epistemic questions, which are treated only in a heuristic way in *Basic Laws*.

Frege took the notion of truth as a normative primitive.[3] He did not leave it unexplicated, and his explications are, as we shall see, highly controversial and involved in metaphysical commitments. But his basic procedure is that of a good scientist in the broadest sense of the term. He created and worked within a theory whose interpretation, *for the fundamental purposes of the science*, was largely uncontroversial. Controversial views were isolated and confined—to the science's heuristic preliminaries and to its frontiers (the philosophical explication of the notions of truth, sense, and assertion, and the application of the logic to intensional contexts, respectively). Extensional logic, more or less as Frege interpreted it, remains fundamental at least in the sense that it is common ground to all logicians and in the sense that its interpretation expresses, with a minimum of controversial accessories, that notion of logical consequence in terms of truth which has traditionally been seen as the central concept of the discipline.

[3] In calling Frege's notion of truth 'normative', I am glossing over a very interesting set of views that he held regarding the normative and descriptive aspects of logic. From the beginning to the end of his career, Frege regarded logic as being descriptive of the laws of logical objects, in particular those of truth. (*PW* 3/*NS* 3; 'Logik' (zwischen 1879 and 1891); Kl 507–8/*KS* 342–3; 'Der Gedanke,' O 58–59.) In fact, Frege seems to have believed that in a sense logic was fundamentally 'descriptive', fundamentally a science of 'being'. Normative restrictions on assertion and judgment derived from 'the way things are' regarding the laws of truth. (Kl 508/*KS* 342; 'Der Gedanke', O 58–59.) To many this view of logic will seem quaint at best. I think that stripped of the particular metaphysics with which Frege endowed it, and supplemented by Quinean and other considerations, it can be made very powerful. In emphasizing that truth is a normative notion, I am not ignoring the 'descriptive' elements in his view. I am simply highlighting a feature of Frege's methodology. Frege attempts to arrive at the laws of truth not by invoking metaphysical assumptions but by concentrating on our practices of assertion, judgment, and deductive inference and by developing his science of logic through reflecting on the 'oughts' of good intellectual practice.

None of this is to deny that Frege had a controversial metaphysics. His philosophical views about truth (particularly Theses (c) and (d) and the 'redundancy' conception), his theory of sense, and his theory of judgment and assertion are widely doubted. Indeed, one might safely count them mistaken.

Frege's philosophical views are not, as such, a set of unfortunate superfluities. I think that a metaphysics—or rather a set of controversial philosophical proposals—in this area can hardly be avoided. There are philosophical questions about truth, meaning, cognitive value, and judgment that are genuinely difficult and apparently genuine. Frege responded to—in fact, in some cases introduced—these questions. And in order to deal with problems about informativeness, about the commitments of propositional-attitude discourse, about the mechanisms of word denotation, and so forth, he postulated certain metaphysical entities (senses, *Gedanken*) that are no less controversial than Russell's facts or propositions. Russell's own theory is in part an attempt to answer these same questions. So from a certain philosophical standpoint it may seem that until these issues are thrashed through, Frege's position holds no advantages over Russell's. His extensional logic owes debts that must be paid before a balance sheet can be drawn up.

There is surely something to this standpoint. But I think that it overlooks one of Frege's central insights and ignores the cognitive advantages of his pragmatic method. Frege's insight is that the normative notion of truth is the central semantical feature of sentences and the fundamental concept of the science. And his pragmatic method of isolating controversy carries the subject a long way before philosophical issues intrude. That has been the method of all successful sciences, mathematical logic included; and success is perhaps our surest guide to knowledge. These considerations tend to favor Frege's basic approach to logic unless the philosophical issues with which he and Russell grappled were decisively and 'scientifically' decided in a way that undermined Frege's extensional starting point. That possibility seems remote.

The Church–Gödel Argument

Here is perhaps a good place to enter into a digression on the relation between Frege's argument for Thesis (b) and an argument proposed by Church and Gödel that is clearly inspired by Frege (Alonzo Church, 'Carnap's *Introduction to Semantics*', *The Philosophical Review*, 52 (1943), 298–305; Kurt Gödel, 'Russell's Mathematical Logic', in P. Benacerraf and H. Putnam (eds.), *Philosophy of Mathematics* (Englewood Cliffs, NJ: Prentice-Hall, 1964).) The argument has a number of interesting variants, and it has been put to even more uses. Gödel's version is particularly rich in implications. I shall, however, discuss only what has become a standardized form.

The argument is supposed to show that all true sentences denote the same thing; an analogous one would show that all false sentences denote the same thing. The argument *first* presupposes that sentences have a semantical feature that bears enough of an analogy to the central semantical feature of terms to be given the same expression. (This presupposition is often not made explicit. For convenience we shall, inaccurately, call it a 'premise'.) Let us dub this feature 'denotation' in accord with Frege's Thesis (a). *Second*, the argument assumes the Composition Principle (1). And *third*, it assumes that logically equivalent expressions have the same denotation. Take any true sentences S and S'; S is logically equivalent with a sentence of the form '$(\iota x)(x = 0) = (\iota x)(x = 0 \,\&\, S)$'. So by the third premise, S and this sentence have the same denotation. But the latter sentence yields the sentence '$(\iota x)(x = 0) = (\iota x)(x = 0 \,\&\, S')$' by substitution of codenotational terms on the right side of the identity sign. So *these* two sentences have the same denotation, by the second premise. But the new sentence is logically equivalent with S'. So by the third premise, they have the same denotation. So S and S' have the same denotation.

Frege accepted not only the conclusion of the argument, but all three premises. But in arguing for the conclusion, in effect Thesis (b), he did not advance this argument. I do not find it plausible to view Frege as giving an elliptical or invalid approximation to this argument. The primary reason for this is the one I have already proposed. Frege invokes the normative foundations of logic and the normative roots of the primacy of sentences in logical theory (and in everyday language use) in arguing for his conclusion. That is, he has a premise about the point of logic; and he connects the notion of sentence denotation both with this point and with his primary analytical tool, the Composition Principle (1). The Church–Gödel argument makes no such appeal to the purpose of logic or semantics.

Another reason why Frege's argument is different can be developed by looking ahead in our discussion. One source of plausibility for the third premise of the Church–Gödel argument derives from the comparison of sentences to terms—in effect, Frege's Thesis (c). Clearly, the denotations of logically equivalent terms are the same. Insofar as sentences are terms, or at least designators, they plausibly fall under the same principle. But many of the considerations that led Frege to accept Thesis (c) presuppose a prior commitment to a semantical analysis of sentences in terms of their truth-values.

A deeper version of the same sort of point can be made from another angle. It may seem perfectly reasonable to accept the third premise of the Church–Gödel argument independently of comparison between the semantics of terms and the semantics of sentences. Suppose that we avoid relying on the view that sentences, like terms, designate or denote entities. The notion of the denotation of a sentence was initially introduced as that notion which captured the primary semantical feature of sentences for logical theory that could be linked up, by the Composition Principle (1), with the denotations of terms. Should we not expect, virtually apriori, that logical theory ought to count sentences as being the same with respect to their primary semantical feature if it counts them logically equivalent?

The rhetorical question packs a punch. But it still overlooks how fundamental Frege's starting point is. The sentences that are indicated to be logically equivalent in the Church–Gödel argument are so counted under a prior conception of logical equivalence, whether informal or fully articulated. This notion already employs some concept of truth—truth under all interpretations, necessary truth, or the like. From Frege's standpoint, this notion of logical consequence (and logical equivalence) already brings with it a commitment to truth-values as the central, logically relevant feature of sentences. So again the third premise of the Church–Gödel argument is less fundamental from Frege's standpoint than its conclusion. Frege's syntactical analysis—Thesis (c)—his conception of logical consequence, and the metaphysics of his logical theory, e.g. Thesis (d), all depend on his commitment to logic's being primarily concerned with the normative notion of truth.

It would be absurd, of course, to suggest that Frege's conception of logical consequence in terms of (necessary) truth was somehow arbitrary, or merely one of many equally suitable choices. The conception lies in the mainline tradition of logic that stretches back to its beginning. Even those conceptions of logic prior to Frege that allowed metaphysical visions to predominate tended to maintain allegiance to the informal conception of logical consequence from which his theory sprang. There have been in this century a few approaches, self-consciously reacting against the main tradition, that have departed from the standard informal conception of logical consequence, following a metaphysical, or more often an epistemological muse. At this point, such approaches must be regarded as secondary developments.

The preceding discussion is not intended to suggest that the Church–Gödel argument is circular. (The third premise is not equivalent to the conclusion.) The argument is usually given in a context in which people already have the ordinary notion of logical consequence, and in which the notion of a denotation for sentences is open to determination. The usual way of reading the argument is to give it the flavor of, 'If you are willing to concede that there is a notion of sentence denotation that meets these restrictions (those of the argument's second and third premises), I will surprise you with what the denotation of a sentence has to be.' The third premise might be bolstered by the argument I gave above [pp. 94–95]. By contrast, Frege already knew exactly how he wanted to use the notion of sentence denotation; it was restricted by the Composition Principle (1) (second premise); but it had to accord with the primary aim of logic, as it has traditionally been conceived.

Much of the surprise of the Church–Gödel argument derives from implicitly thinking of sentence denotation primarily in terms of the pretheoretical notion of naming, rather than primarily in terms of a specific conception of its theoretical employment, as Frege did. Once one has taken the dubious step of seeing sentences as names, or at least as designating some entities that are functionally dependent on word-denotation, it is intuitively surprising to think of them as designating only one of two entities, and odd to reify truth and falsity. Having come so far, it is perhaps more intuitive to take sentences as 'designating' possible states of affairs, or something like that. At least many have thought so. Seen this way, the conclusion of the Church–Gödel argument is unappealing. It is doubtful, however, whether anyone, except perhaps for Church, has endorsed the argument read in this way.[4] The natural and most common response to the argument is to reject its first 'premise': sentences do not name, refer to, or designate any entity.

As I have indicated, it is possible to see the argument as using a less determinate notion of denotation that gets around this objection. One can consider the argument without adopting Frege's Theses (c) and (d). Then the oddity of the conclusion disappears—the better for reflecting on the logical relationships that the argument reveals.[5]

[4] Even he enters qualifications: Alonzo Church, *Introduction to Mathematical Logic* (Princeton: Princeton University Press, 1956), p. 25 n. 66.

[5] Prior to the argument's formulation, as early as 1903, Russell rejected the first premise. He insisted that the relevant semantical notions for sentences are quite different from those for terms. It is a bit open to question how seriously this rejection is to be taken. Russell refused to call the relation between a sentence and a fact or proposition the relation of 'naming' or 'denoting'. But he treats propositions or facts as complexes made up entirely of entities that have traditionally been thought of as referred to by words—properties, relations, individuals and the like. And he regards these complexes as playing a central role in his semantical theory of sentences. Thus at times Russell's point seems to come to little more than that sentences are not *ordinarily speaking* names, a point with which Frege might well agree. Actually, of course, the issue in Russell is quite complex. (Cf. e.g. the first lecture in 'The Philosophy of Logical Atomism' in R. C. Marsh (ed.), *Logic and Knowledge: Essays 1901–1950* (London: Unwin Hyman, 1956).)

Russell remains at odds with the Church–Gödel argument even in its less titillating form—even after sentence denotation is construed as the yet-to-be-determined semantical feature that is connected to the denotation of terms by means of the Composition Principle (1). As I mentioned earlier, Russell may be interpreted as rejecting the third premise of the argument. He accepted more or less the traditional conception of logical equivalence and judged logical consequence in terms of the traditional modalized notion of truth (e.g. *PM* section $A^{*}1$). But he took the primary semantical correlates of sentences to be what he called 'propositions' and sometimes 'facts'. If one interprets Russell as accepting the first 'premise' of the Church–Gödel argument by granting it the liberal conception of sentence denotation that involves no commitment to Thesis (c) (so sentence denotation is merely the central semantical feature of sentences in one's formal semantics), then one must see Russell as rejecting the argument's third premise. For then facts or 'propositions' are sentence denotations; and logically equivalent facts could, on Russell's view, differ.

What allows this position to remain compatible with the principle that exchange of co-denotational expressions preserves truth-value is, of course, the theory of descriptions. This theory by itself blocks the Church–Gödel argument by depriving one's language (artifically, I think) of the definite description operator, or any comparable device for forming complex singular terms that have denotations.

A discussion of the Church–Gödel argument that is Russellian in its metaphysical cast occurs in Jon Barwise and John Perry, 'Semantic Innocence and Uncompromising Situations', *Midwest Studies*, 6 (1981). Unfortunately, the paper contains much that is misleading. Frege's arguments are dismissed in two paragraphs. One paragraph characterizes Frege's rhetorical question 'What else besides truth-value is compatible with the composition principle?' as a metaphysical oversight. This dismissal would perhaps be fair if it did not ignore Frege's normative motivations and methodology. Frege's argument from our primary interest in sentences, glossed in one sentence, is countered by an irrelevant appeal to embedded sentences (irrelevant because the reason sentences are interesting is that they are the vehicles of assertion and judgment). The Church–Gödel argument is discussed only on the

Theses (c) and (d)—Pragmatic Motivations

In my view, the first of Frege's arguments for (b) and both his arguments for (a) are sound. Although (a) and (b) have often been targets of criticism, most of the criticism stems from construing (a) and (b) in the light of Theses (c) and (d). I believe that doubts about (c) and (d) are justified. But as I shall try to show in the remainder of this section and in Section II, such doubts are less interesting than has sometimes been supposed. The discussion of these theses will serve to introduce background essential to Frege's treatment of logical objects.[6]

In a letter to Frege in 1903, Russell challenged Thesis (c), the view that sentences are to be regarded as of the same logical type as singular terms: 'I have read your essay on sense and denotation, but I am still in doubt over your theory of truth-values, only because it seems paradoxical to me. I believe that a judgment, or even a thought, is something so completely peculiar that the theory of proper names has no application to it' (*PMC* 155–6/*WB* 238; Russell to Frege, 2/20/1903). The gist of Russell's challenge has been repeated by subsequent generations, and with qualifications to hedge against the overstatement in Russell's phrase "no application", I would echo it. But it is easy to be led by the paradoxical ring of (c) and (d), as I think Russell was led, into misunderstanding their import and place in Frege's system.

Frege repeatedly emphasized intra-logical, pragmatic advantages for regarding truth-values as objects: 'How much simpler and sharper every thing becomes through the introduction of truth-values, only thorough occupation with this book can show. These advantages alone already put a great weight on the balance in favor of my conception, which indeed may seem

unquestioned (but widely rejected) assumption that sentences as wholes 'designate' some entity. And it is resisted as if it had been presented on this unquestioned assumption, and widely accepted, as a 'proof' from 'virtually apriori' first principles. The incompatibility of Russell's system with the argument's conclusion has been widely recognized. No one has tried to utilize the argument to refute Russell, least of all Church and Gödel, who were, of course, thoroughly familiar with Russell's system. The role of the argument in the history of semantics is more subtle than treating it as a proof from purportedly obvious first principles could suggest.

[6] The relation between Theses (c) and (d) is a complicated and subtle matter. Ontologically, of course, logical objects, such as truth-values, and functions on these objects are prior for Frege to singular terms and function signs. They exist before the signs existed, and would have existed regardless of whether the signs did. On the other hand, Frege held, on several occasions, that one could not engage in reasonably sophisticated thought except by means of language. The analysis of thought relied heavily on analysis of linguistic structure. But even in the analysis of thought, the analysis of language was not prior in any simple sense. For thinking could and did correct the deficiencies of language. In the light of all this, there is no simple answer to the question of whether Frege reasoned from Thesis (c) to Thesis (d) or vice-versa. Part of why he concluded that numbers are objects and numerals are terms was that he was able to give explicit definitions which amounted to a criterion of identity for the numbers. On the other hand, much of his reasoning to this conclusion was based on observations regarding the structure of mathematical language. Similarly, his reasoning about Theses (c) and (d) is a mixture of considerations regarding the role of objects in logic and the anatomy of the language of logic (properly construed). I shall therefore treat Theses (c) and (d) more or less together, without trying to sort out the various relations of relative priority that obtain between them.

strange at first glance' (*BL* 7/*GG* x). In fact, *The Basic Laws of Arithmetic* mentions only considerations involving simplification of logical theory as motivations for (c) and (d). These considerations are also dominant in Frege's post-paradox writing.

I should make it clear here that in calling Frege's reasoning 'pragmatic' or 'intra-logical', I am not suggesting that he took the commitments that he based on such reasoning to be less than absolutely serious. Such commitments were not merely practical conveniences or technical artifices. Frege saw himself as making objective discoveries. What I wish to emphasize is the great extent to which Frege tried to develop his positions from his analysis of logical structure and from observations regarding functional analogies between different components of that structure. In his arguments for (a)–(d), considerations that derive from intuitions not firmly entrenched in the actual practice of logic, 'metaphysical intuitions', play a secondary role in Frege's argumentation. Whatever conceptions most profoundly clarified and simplified logical theory, whatever language made mathematical practice more rigorous, more comprehensive, more fruitful, and less *ad hoc*, were seen as providing insight into the most abstract features of the world.

The pragmatic cast of Frege's thought seems to have come naturally. Only rarely did he remark on his methodology in general terms. The effusion from *Basic Laws*, quoted in the previous paragraph but one, constitutes a relatively unusual example. The following passage from a thrice rejected manuscript of 1880–1 provides another:

> All these [mathematical] concepts have been developed in science [Frege terms mathematics a science] and have proved themselves fruitful. What we can perceive in them therefore has a far higher claim on our attention than anything that everyday trains of thought might offer. For fruitfulness is the touchstone of concepts, and the scientific workshop is the real field of observation for logic. (*PW* 33/*NS* 36–7, 'Booles rechnende Logik und die Begriffsschrift' (1880/1881); cf. also *KS* 124, O 161, 'Über das Trägheitsgesetz'; *KS* 369, O 150; 'Die Verneinung')

Frege's pragmatic considerations rest on analogies that are quite natural within a formal context. In formulating the propositional calculus, it is natural to quantify the letters that stand for sentences in something like the way one quantifies into the places held by singular terms in the first-order functional calculus. The places for sentences, like the places for terms, stand in argument places for functional expressions; they sometimes constitute value expressions resulting from functional application; and they never stand for functional or predicate expressions. When one quantifies the letters that stand for sentences, the natural interpretation of the domain of quantification is to take it as consisting of the two truth-values—as several generations of logic students have been made aware (Cf. *PMC* 158/*WB* 241; Frege to Russell, 5/21/1903; *KS* 225–6; O 368–369.)[7]

[7] There are variants of the analogy between terms and sentences as regards their sense and denotation that Frege mentions, but which I shall skim over. For example, he cites his theory of

Another analogy, which is more debatable, is that between nondenoting names and truth-valueless sentences in natural language (G & B, 'S & R' 62/*KS* 148; O 62; *PMC* 152/*WB* 235; Frege to Russell, 12/28/1902). Both maintain a sense in the absence of denotation. Frege thought that subject–predicate sentences containing nondenoting terms always lacked truth-values.

It is difficult to see to what extent he accepted this view on intuitive grounds and to what extent he reasoned to it. If one already sees predicates as denoting functions, then one will see a nondenoting name as providing no argument for such a function. Functions without argument yield no value. And if the denotations of sentences are the values of such functions, and are counted truth-values, then sentences involving only the application of predicates to nondenoting terms (and application of functors to sentences so obtained) will lack truth-value. This reasoning, of course, assumes the assimilation of predicates to function signs. And this assimilation is tantamount, as we shall see, to accepting Thesis (c). So the reasoning cannot be seen as providing much independent support for Thesis (c).

Very likely, the view that nondenoting terms in subject–predicate sentences yield truth-valueless sentences was also found acceptable by Frege on intuitive grounds. The various examples he gives do elicit in many the intuition that the sentences are neither true nor false. But there are numerous other cases that are at best indecisive witnesses for Frege's defense. Since this issue has been discussed at uncommon length by others, I shall not go into it. I think that Frege's view of the intuitive relation between nondenoting terms and truth-valueless sentences was not very critical to his account of truth-values. Since he banned nondenoting terms from his formal theory, he rested little weight on the point.

An analogy of which Frege makes more is that between nonassertive occurrences of declarative sentences (suppositions or occurrences within other sentences) and proper names (*BL* 35/*GG* 7). Sentences often occur embedded in other sentences (for example, in the antecedents of conditionals) in such a way as to contribute to semantical structure, without being asserted—like terms. Moreover, whole sentences can be put forward merely for consideration without carrying any assertive force—again like terms.

These analogies between sentences and terms are, of course, not very gripping. They take on interest when linked to Frege's larger strategy. One of Frege's most profound contributions was to separate the notions of predication and assertion. More generally, he distinguished the notions of logical structure and pragmatically relevant force. The deeper point of the present analogies is that within a formal theory that attempts to lay bare semantical

indirect discourse as tending to confirm the introduction of truth-values (*BL* 7/*GG* x). The idea seems to be that, in indirect discourse, just as terms shift from denoting their customary denotations to denoting their customary senses, so sentences shift from denoting their truth-values to denoting the thoughts they customarily express. On Frege's conception, there need be no shift in grammatical category between the occurrence of a sentence standing alone, and its occurrance in indirect discourse, which is clearly a singular term position.

structure, one can prescind from the primary difference between names and sentences (that only the latter can be used to effect linguistic acts or thoughts, prototypically assertions and judgments). The difference between names and sentences can be taken to lie in their point, their use, not in the form of their contribution to semantical structure. Actually, as we shall see in Section II, Frege's formal theory did make formal distinctions between sentences and terms. But the distinctions do not leap to the eye. Although one might believe (as I do) that form should correspond more closely to use than Frege's logical theory allows, subsequent formal usage has confirmed that Frege's analogy constitutes an insight that affords at least a convenient alternative in setting up a logical system.

Frege's construal of predicates as functional expressions is perhaps the most obvious and widely appreciated ground for Theses (c) and (d) and for his view that truth-values are objects. As far back as the *Begriffsschrift* in 1879, Frege had interpreted predicates as function signs (*B*, section 9). Once he supplemented this initial conception with an explicit semantical analysis, which he arrived at by the early 1890s at latest, he was forced to think of functions as denotations for predicates. He called such denotations 'concepts'. (We shall limit considerations to 1st-level concepts.) Objects, obviously, served as arguments for (1st-level) concepts. But then there must be values for these functions. These must be the denotations of sentences. Sentences are not themselves functional expressions, so their denotations are not functions. Moreover, the values of prototypical functions, the denotations of prototypical completions of functional expressions (terms), just are objects. (We shall, for now, regard an object as anything that is denoted by what is, under logical analysis, a term.) Taking concepts literally to be functions was tantamount to taking the denotations of sentences to be objects (and the completions of predicates, sentences, to be terms). Since Frege had independent grounds for regarding the denotations of sentences to be truth-values, this line of thought entailed that truth-values were objects.[8]

Why did Frege take concepts, denotations of predicates, *literally* to be functions? One primary and lasting motivation was, yet again, pragmatic. Seeing sentences as created by the application of functional expressions effected a simplification in the understanding of the composition principles. The simplest construal of the Composition Principle (1) is to take 'the

[8] One might be tempted to think that Thesis (a), or at any rate (b), already commits Frege to taking truth-values to be objects. For by our stipulation, an object is anything that is denoted by a term. But the phrase 'the denotation of Sentence 5' (cf. Thesis (b)) is a term. So by Thesis (b), truth-values are objects. This line of reasoning misses the fact that the phrase 'the denotation of a sentence' need not be a term *under logical analysis*. Similarly, for the phrase that begins Principle (1). 'Denotation' as applied to sentences in the initial construal of Thesis (b) is guided by the compositional method, *loosely* expressed in Principle (1). As far as Frege's arguments for (a) and (b) are concerned, there need be no entity that could be called under logical analysis 'the denotation' of the sentence. Of course, once Frege has committed himself to Theses (c) and (d), he can consider Theses (a) and (b) to have proper construals, under logical analysis, that commit him to taking truth-values to be objects.

denotation of sentence *s*' to be a singular term, denoting an object. (Cf. Note 8.)

A closely related motive was to provide a simple formal expression of the formal analogies between predicates and function signs (*BL* 6, 34–5/*GG* x, 6–7; *PW* 235, 243–4/*NS* 253–4, 263; 'Logik in der Mathematik' (1914)). Like function signs, predicates have open places for terms. The primary role of predicates from the point of view of logic is functional—to take objects into truth-values.

Completeness and Incompleteness

Frege's pragmatic motives are, I think, dominant. But the analogy between predicates and function signs is sometimes associated by Frege with remarks that have a darker, more metaphysical hue—remarks about similarity in their 'unsaturatedness' or 'incompleteness'. He says that the essence of a function is its making a connection between its arguments and its values, in a specific sort of 'need for completion' (*BL* 33–4/*GG* 5–6; G & B, 'C & O', 47/*KS* 171; O 197–198; G & B, 'F & C' 24–5/*KS* 128–9; O 7–8). Moreover, he writes, 'An object is anything that is not a function, so that an expression for it does not contain any empty place. A declarative sentence contains no empty place and on that account its denotation is to be regarded as an object' (*G & B*, 'F & C' 32/*K*S 134; O 18).

Here Frege may appear to be inferring from a metaphysical thesis about incompleteness of functions and completeness of objects and from a thesis about how language must match reality as regards completeness or incompleteness, to the conclusion that truth-values are objects. Michael Dummett interprets the characterization of objects as anything that is not a function in this passage (and in an equivalent one in *Basic Laws*, vol. I, section 2) as an *ad hoc* attempt to induce the reader to accept truth-values as objects (*IFP* 235n). Neither the metaphysical reading nor Dummett's attribution of desperate improvisation places Frege in a very attractive light.

To begin with the latter interpretation, I do not find Dummett's charge plausible. Frege's characterization of objects is independent of Thesis (d) and precedes its adoption. The idea that objects can be recognized as whatever is never the denotation of an incomplete, functional expression goes back at least to *The Foundations of Arithmetic*. (Cf. pp. x, 77n, 72—before Frege's adoption of Thesis (d).) In the latter passage, *FA* 72, Frege writes that the point of counting number words as words for objects (or self-subsistent objects) is 'only to preclude the use of such words as predicates or attributes . . . ' (Cf. *PW* 100, 104–105/*NS* 109, 113–114; 'Über den Begriff der Zahl: Auseinandersetzung mit Kerry' (1891–1892).) Given Frege's view that truth-values are denotations of complete sentences, and never denotations of predicates, and given this characterization, truth-values fill the bill as objects.

Frege's characterization of an object as the denotation of any expression other than a predicate or function sign may seem either to emasculate the notion of object or (perhaps equivalently) to commit one to objects too easily. In discussing Frege's arguments for Thesis (a) we attributed to him a notion of sentence denotation that does not carry genuine ontological commitment. But now, it may seem, we are allowing Frege to smuggle ontological commitments into his arguments for Theses (a) and (b) by granting him an excessively liberal criterion for ontological commitment to objects. I have argued that the relevant criterion was not fabricated, as Dummett suggests, simply to make palatable the view that truth-values are objects. But it may seem that Frege made illegitimate use of a criterion that was first developed in a context in which the denotation of sentences was not an ontological issue—resorting to a cheap means for ontological gain.

There is something to this worry. I believe, however, that it cannot be taken at face value. Frege does argue from his characterization of objects to Thesis (d) (in the paragraph following the relevant characterization of objects in 'Function and Concept'). But he does not take the characterization as stipulated or ungrounded. In the first place, there are substantial analogical considerations that underly his counting sentences and terms 'complete' and predicates and function signs 'incomplete'. In the second place, Frege seems to have always regarded the characterization of objects as resting on an antecedent notion of completeness that he believed he could apply to sentences (and truth-values) as well as to terms and ordinary objects. It is the notion of completeness that bears the weight, not the bare claim that objects are the denotations of every sort of symbol other than function signs. The intuitive notion of completeness underlies and motivates the syntactical, semantical, and ontological claims.

Then isn't Frege's engrossment in the completeness–incompleteness distinction simply a metaphysical indulgence? I would not deny that some of Frege's uses of the distinction involve a kind of fixation that is difficult to fathom, much less defend. But his deployment of the distinction to support his view that the denotations of sentences, truth-values, are objects, seems to me less problematic than some other uses he makes of the distinction.

Let us lay aside Frege's view that no objects are functions and no functions are objects. I think that this view is extremely doubtful and that it probably does constitute an instance in which Frege allowed his sound conceptions of logical function to harden unnecessarily into a metaphysical doctrine. These matters are, however, intertwined with a suprisingly large number of serious considerations (for some of them, see Burge 'Frege on Extensions of Concepts, From 1884 to 1903', *The Philosophical Review*, 93 (1984), 3–34; Chapter 7 below). I shall avoid the tangle here.

Let us consider only Frege's views that in using (what were under logical analysis) function signs, one is committed to their denotations, functions; and that in using (what were under logical analysis) terms, one is committed to their denotations, objects. As we have noted, predicates are like function signs in having empty argument places, and in having a functional role in logical theory. Sentences are like terms in not manifesting such formal incompleteness and in not having a functional role. In concluding that the denotations of sentences are objects, Frege may be reasonably seen *not* as drawing a primitively minded inference from some pre-Socratic vision of the world as a mixture of the complete and incomplete—but as simply summing up and embellishing the analogies, within his logical system, between the roles of sentences and terms, and their contrasts with predicates and function signs.

The mapping of objects and functions onto truth-values—the central semantical feature of sentences—is *the* primary formal role of predicate expressions (or concepts) within formal logical theory. The deep differences between predicates and ordinary function signs, and between sentences and terms, were largely shunted off into the theory of force or use. Frege did not lose sight of the differences. But he thought that he could draw ontological conclusions from a semantical theory that abstracted from them. In regarding concepts as functions and truth-values as objects on grounds of the 'incompleteness' of signs for the latter, Frege was basing ontological commitments on the semantical analysis of the logical forms of sentences in whose truth he believed. Frege's methods, if not his conclusions, seem unexceptionable.

Clarification of Extensions of Concepts

The assimilation of concepts to functions served one other large purpose in Frege's system. It provided the key to his attempt to clarify the notions of a concept and of the extension of a concept. As I have tried to show in some detail elsewhere, Frege was unclear about and dissatisfied with these notions from the time he first introduced the latter in *Foundations* (section 68) (1884) up to and through the publication of *Basic Laws* (1903). (Cf. Burge, 'Frege on Extensions of Concepts' (Chapter 7 below).) The key to the clarification that he attempted, until Russell unsettled him, was the notion of the course of values of a function. Frege sought to make this notion intuitive by appeal to the graph of a function (G & B, 'F & C' 25/*KS* 129; O 8), which he seemed to think of both algebraically and geometrically. He self-consciously did not interpret the graphs as sets of ordered pairs for a variety of reasons deriving from his emphasis on the priority of functions over their courses of values. We shall return to these points in Section III.

The notion of a concept had had a long but mathematically barren career in the logical tradition. It was not held in high esteem by mathematicians in Frege's day. By contrast, the notion of a function was well established in mathematics. By assimilating the denotations of predicates to those of function signs—giving them a recognizable mathematical role—Frege hoped to clarify the notion of a concept and burnish its reputation. At the same time, he would be effecting a unification of the languages of logic and mathematics in accord with his logicist thesis. This motivation is explicit when Frege first introduces the assimilation of concepts to functions:

> ... for what purpose, then, are the signs '=', '>', '<' admitted into the circle of those that help form a functional expression? It seems that nowadays more and more supporters are being won to the view that arithmetic is further-developed logic ... I too am of this opinion, and I base upon it the requirement that the symbolic language of arithmetic must be expanded into logical symbolism. (G & B, 'F & C', 30/*KS* 132; O 15)

The clarification of the notion of a concept was intended to give a firm foundation to those objects logically associated with concepts (these 'extensions') with which in *Foundations*, driven by grammatical considerations and his logicist goal, Frege wanted to identify the numbers. The extension of a concept was understood as the course of values or graph, obtained by providing all objects one by one as arguments for the concept (function) and taking the resulting truth-values as values. The whole procedure, taken as a completed whole, was what Frege regarded as a logical object. Since the introduction of such objects crucially depended on Axiom V, which led to Russell's paradox, Frege's attempt to clarify the notion of the extension of a concept by assimilating concepts to functions failed.

The Redundancy Conception of Truth and the Notion of Object

I shall conclude our discussion of Frege's reasons for accepting Theses (c) and (d) by considering his redundancy view of truth. In 'On Sense and Denotation' he writes:

> One might be tempted to regard the relation of the thought to the True not as that of sense to denotation but rather as that of subject to predicate. One can, indeed, say: "The thought that 5 is a prime number, is true." But if one observes more closely, one notices that really nothing more is thereby said than in the sentence '5 is a prime number.' (G & B, 'S & R' 64/*KS* 150; O 34). (Cf. also *PW* 128–9, 233–4, 251–2, 255–6/*NS* 139–140; 'Logik' (1897); *NS* 251–2; 'Logik in der Mathematik' (1914); *NS* 271–2; 'Meine grundlegende logischen Einsichten' (1915); *NS* 275–6; 'Aufzeichnungen für Ludwig Darmstaedter' (1919); Kl 514/*KS* 347; O 63; 'Der Gedanke' (1918).)

Frege uses the view to ward off possible doubts about the postulation of the truth-values as objects denoted by all sentences, regardless of subject matter. If truth were an attribute of a limited range of entities (thoughts), it would be difficult to motivate the claim that every sentence denotes one of the truth-values and his view that (in a sense to be sharpened in Section III) all assertive uses of sentences regardless of subject matter are committed to the object truth.

In 'On Sense and Denotation', two large philosophical ideas emerge in connection with the redundancy conception. One utilizes truth-values as objects in an account of assertion and judgment. The other bears on scepticism. We shall consider these themes in turn.

Frege goes on from the passage just cited to argue that the claim or judgment that a thought is true arises not from the predication of 'is true' of a thought, since the sentences 'the thought that 5 is prime is true' and '5 is prime' express the same thought regardless of whether they are used with or without assertive force. Truth claims or judgments depend on the combination of the form of a declarative sentence with its 'usual force'. Frege thinks that such truth claims are indicative of the real relation between a sentence or thought and its truth-value (G & B 'S & R' 65/*KS* 150; O 34–35). Judgments are 'advances from thoughts to truth-values'. Since truth claims and judgments cannot be represented in subject-predicate form, the relation of a sentence or thought to its truth-value cannot be regarded as that of subsumption of a thought (sentence) under a property. Frege proposes that the appropriate relation is that of a sentence or thought to its denotation. (Cf. also *PW* 128–9, 233–4, 251–2/*NS* 139–40; 'Logik' (1897); *NS* 252–3; 'Logik in der Mathematik' (1914); *NS*, 271–2; 'Meine grundlegenden logischen Einsichten' (1915).)

One need hardly note that considered as an argument for Thesis (d), this is pretty weak. (It is doubtful that Frege intended it as such.) One could respond that on Frege's own account, two sentences could have the same assertive force—both could count as assertions—while one lacked a truth-value and the other had one. So truth-values as objects cannot be essential to the account of assertive force. Even if this reply is not decisive, Frege does not show why it is not.

I think Frege was here again thinking analogically. Normally, the point of using names was to secure a denotation, a bearer, to relate a mode of presentation to an object. Normally, the point of using sentences, what 'matters to us', is to claim truth for a thought. The object, in the sense of the point or *objective*, of sentence use is truth. It is illuminating therefore to see truth as an object. There is more than a suggestion of this reasoning when Frege writes:

> The designation of truth-values as objects may here appear as arbitrary fancy or perhaps a mere play on words, out of which no profound consequences could be drawn. What I call an object can be more exactly articulated only in connection with concept and relation. . . . But so much should already be clear, that in every judgment, no matter how trivial, the step from the level of thoughts to the level of denotations (the objective) has already been taken. (G & B, 'S & R' 63–4/*KS* 149; O 34)

The parenthetical phrase is the key to the passage.

To many this reasoning may seem indeed to rely on a mere pun on the word 'object'. I think that there is more to it than that. Both the relevant objective of sentence use, truth, and objects that are denoted by terms are for Frege mind-independent. And in some sense they are what sentences and terms are respectively 'about'. (In fact, as Frege emphasizes in his arguments for (a) and (b), objects are of interest to us because and only because of the objective of assertion and judgment.) Both points resist simple or quick put-downs. Both mind independence and being the topic of a discourse are involved in traditional explications, stemming from Aristotle, of the notion of object.

I am not suggesting that the analogies between the 'objects' of terms and the 'object' (objective) of sentence use provide a sound argument for Frege's assimilation. I believe the contrary. In fact, I believe not only that the analogies are not compelling, but also that Frege's redundancy view of truth, which motivates them, is untenable. (Part of the reason for this untenability lies in the semantical paradoxes; cf. Section III.) Rather, what I am suggesting is how Frege might have come to see the analogy as intuitively attractive, given his view that the attribution of truth adds nothing to a thought. We shall further articulate Frege's analogy between objects and the objective of assertion, in Section III.

Frege puts Thesis (d) and the redundancy view of truth to use as a weapon against the sceptic about an objective world. Frege writes that the True and the False 'are recognized, if only implicitly, by everybody who judges something to be true—and so even by the sceptic' (G & B, 'S & R' 63/*KS* 149; O 34). (Frege assumes contrary to the legends about Pyrrho, but probably correctly, that no sceptic suspends all judgments.) The idea is that *every* act of judgment aims at truth and presupposes some discrimination between truth and falsity. Frege explicates the point by his redundancy thesis: truth-values are not a property of thought, where thoughts constitute one subject matter among many: 'What distinguishes [truth] from all other predicates is that it is always asserted when anything at all is asserted' (*PW* 129/*NS* 140; 'Logik' (1897)). Since the true is an object logically associated with the truth predicate and so with judgment—a logical, mind-independent object—judgment itself presupposes an objective world. We shall sharpen Frege's point in Section III. I think that one could probably dispense with the implausibilities of the redundancy view to provide the sort of premise needed for joining with (d) in order to defeat the relevant sceptic. If only (d) were true!

Gödel remarked that Frege held the view that all true sentences have the same denotation 'in an almost metaphysical sense' (Gödel, 'Russell's Mathematical Logic', 214). It is true that Frege puts the doctrine to use against the sceptic. There is no question but that he thought of truth as an object. And there are some unfortunate, but qualified and never repeated remarks in 'On Sense and Denotation' about parts of the True (G & B, 'S & R' 65/*KS* 150–1; O 35)—remarks that prompted Gödel (inappropriately, I think) to compare Frege's view with Parmenides'. But if the reasons for a view may be seen as an index of its character, Frege's doctrine cannot comfortably be called metaphysical. The brunt of his case for (d) rests on the formal simplifications the view effects within his logical theory and the clarification it was supposed to yield for the notion of the extension of a concept. With the discovery of Russell's paradox, the latter support was undermined, leaving only the former.

It is interesting that in his post-paradox period, Frege cites only considerations of simplification (for example, the congeniality of the view with the Composition Principle (1)) in favor of Thesis (d). In his epistolary responses to Russell's doubts he remains doggedly within the elegant confines of his logical theory—repeatedly employing the Composition Principle and pointing out difficulties in Russell's vague but seminal alternatives. In 1906 at the beginning of the scrap 'What May I Regard as the Result of my Work?', he cites 'a concept construed as a function' and introduces the citation with the remark, 'It is almost all tied up with the *Begriffsschrift*' (*PW* 184/*NS* 200; 'Was kann ich als Ergebnis meiner Arbeit ansehen?' (1906)). In his late writings he gives up on the notion of the extension of a concept, and in 'The Thought' (1918) the arguments against scepticism make no use of truth-values.

Once Frege's intra-logical analogies are appreciated, there is, I think, no need other than momentary expositional convenience to treat sentences as of the same logical type as names. One may maintain in one's semantical theory a reflection of the large differences in use between sentences and terms. And one may return to the natural view that terms, not including sentences, are the basic avenue of ontological commitment.

The primary reason why Frege did not take this more modern view of the matter in his great, pre-paradox writings is that he wanted to use the truth-values in his account of logical objects. Logical objects were needed for his logicist project—the project of showing that the mathematics of number is reducible to logic. For mathematics was apparently committed to objects, the numbers; and to account for these commitments Frege thought he had to generate commitments to appropriate sorts of objects within logic. The truth-values were the basic logical objects from which all others were to be generated. (Cf. Section III.)

By roughly 1906, however, Frege seems to have given up logicism. So the most prominent philosophical motivation for postulating logical objects lapsed. The doctrine that truth-values are objects may have become less important to him in his later years. He does not give up the view, however. And I suspect that in addition to analogical or pragmatic considerations, he retained a philosophical motive for holding it. This motive was his desire to explicate the objectivity and informativeness of logic—its 'descriptive' as well as normative character. (Cf. note 3 and Section III.) Although I shall not discuss it here, I think that this motive is profoundly conceived. But after the failure of Frege's logicist project, the attempt to utilize the truth-values as means to articulate the motive was deprived of a coherent background theory within which to bring together a conception of truth with a conception of logical objects. So Frege is left without a theory within which he could argue for using Thesis (d) to articulate his thoroughly unKantian view that logic is an informative science of 'being' (Kl 508/*KS* 342; 'Der Gedanke', O 58).

We have briefly touched on Frege's view about logical objects in our discussions of his attempt to clarify the notion of the extension of a concept, his argument against the sceptic, and his conception of the nature of logic. We shall return to them in Section III. But first, I want to consider an influential body of thought that seems to me to have placed Frege's views on truth-values in the wrong light. This discussion will enable us to develop further Frege's conception of truth, a conception that will dominate our concluding reflections on logical objects.

II

In his two books on Frege, Michael Dummett maintains, as against Theses (c) and (d), that sentences are not names, and truth-values are not objects. As is plain, I do not dispute this conclusion. It is the reasoning behind Dummett's rejection of these theses, and the urgency with which he invests that rejection, that constitute, in my opinion, a serious misrepresentation of Frege.[9]

We have already quoted Dummett's statement that Frege's acceptance of Theses (c) and (d) was an almost unmitigated disaster. For, Dummett writes,

> . . . it obscured the crucial fact that the utterance of a sentence, a complex term . . . can be used to effect a linguistic act, to make an assertion, give a command . . . the general notion of the sense of a word will now have to be taken to consist in the contribution which that word makes to determining what a complex singular term, in which it may occur, stands for, rather than what are the truth-conditions of a sentence in which it may occur. (*FPL* 7)

> If sentences are merely a special case of complex proper names, . . . then, after all, there is nothing unique about sentences: whatever was thought to be special about them should be ascribed, rather, to proper names—complete expressions—in general. This was the most disastrous of the effects of the misbegotten doctrine that sentences are a species of complex name . . . : to rob him of the insight that sentences play a unique role, and that the role of almost every other linguistic expression . . . consists in its part in forming sentences. (*FPL* 196; cf. 643–5)

Dummett takes the adoption of Thesis (c) to underlie the relative inconspicuousness, in Frege's later work, of statements of the context principles, state-

[9] I do think that Dummett neglects to convey the richness and inter-related nature of the theoretical considerations supporting Frege's theses. His remark about Thesis (c) that it is a 'ludicrous deviation' from the forms of natural language and a 'gratuitous blunder' (*FPL* 184) is, at best, immoderate. Incidentally, in *FPL*, Dummett spends two-thirds of his chapter 'Truth Value and Reference' on Frege's view that a sentence with a nondenoting name thereby lacks a truth-value. As I mentioned in Section I, I think that Frege rested little weight on this view in defending Thesis (d). I think Frege regarded his view as a consequence of the rest of his doctrine. Since the consequence accorded with his intuitions, it had some value for him in confirming the doctrine.

ments which had been so prominent in *The Foundations of Arithmetic* (p. x, and sections 60, 62, 106). Dummett's idea is that since Frege assimilated sentences to complex singular terms, he 'debarred himself from a direct statement of the context principle, since this would have involved acknowledging a difference in logical role, of utmost importance, between sentences and proper names of objects other than truth-values' (*IFP* 371). Dummett cites Frege's conclusion to section 10 of *The Basic Laws of Arithmetic* (cf. also sections 29, 31–2) as evidence that only a weakened, generalized analog of the context principle for denotation was still adhered to: The denotations of terms are fixed when it has been determined for every primitive function [whether a concept or not] what the value of the function is to be for the denotations of any terms as argument(s) (*IFP* 408 ff.). In this principle, sentences and predication are given no special prominence over terms and ordinary functional application. Dummett goes on to question the coherence of the resulting doctrine.

Now there is much in Dummett's discussion that we cannot take time to go into. The context principles form an exceedingly complex topic. Despite my disagreement on some fundamental matters in this area, I think that Dummett has contributed a great deal to our understanding of the issues. Here I shall concentrate on disagreements that bear most directly on truth-values.

There is evidence that Frege did not lose sight of the 'crucial fact' that the utterance of a sentence, unlike a term, can be used to make an assertion; that he did not draw the unsound inference Dummett does that 'if sentences are merely a special case of proper names . . . then, after all, there is nothing unique about sentences . . . '; and that Frege was never robbed of the insight 'that sentences play a unique role'.

In the first place, there are a great number of passages throughout his career and especially from the 1890s onward, in which Frege asserts that the aim of logic is to understand the laws of truth (*PW* 2–3, 128–9, 149, 197–8, 252, 253/ *NS* 2–3; 'Logik' (zwischen 1879 and 1891); *NS*, 139–40, 161; 'Logik' (1897); *NS* 212–13; 'Einleitung in die Logik' (1906); *NS*, 272; 'Meine grundlegenden logischen Einsichten' (1915); *NS*, 273; 'Aufzeichnungen für Ludwig Darmstaedter' (1919); Kl 505 ff./*KS* 342 ff; 'Der Gedanke', O 58 ff.) He repeatedly characterizes these laws as normative restrictions on judgment and assertion. Predications of truth are not really distinguishable from the assertoric form of any declarative sentence at all (*PW* 129, 233/*NS* 140; 'Logik' (1897); *NS*, 251; 'Logik in der Mathematik' (1914); cf. our discussion, Section I, of the redundancy theory of truth.) Once he writes that the essence of logic lies in assertoric, or judgmental, force (*PW* 252/*NS* 272; 'Meine grundlegenden logischen Einsichten' (1915)). The vehicle of judgment is a thought and the vehicle of assertion (the expression of a judgment) is a sentence (*PW* 126, 131, 206/*NS* 157, 142; 'Logik' (1897); *NS*, 222–3; 'Logik in der Mathematik' (1914)). Thus the essence and aim of logic is repeatedly associated with sentences and thoughts (the senses of declarative sentences) and their logically relevant uses. The denotations of terms are almost never discussed except in the larger context of this emphasis on the centrality of truth, judgments, thought, assertion, and sentencehood. And in Frege's last years, the denotations of terms receive very little attention at all.

Moreover, there are the passages from the 1890s and later, quoted in Section I, that occur in Frege's arguments for Theses (a) and (b) (G & B, 'S & R' 63/*KS* 149; O 33; *PMC* 152, 158, 163n, 165/*WB* 235; Frege to Russell, 12/28/1902; *WB*, 240; Frege to Russell, 5/21/1903; *WB* 245n; Frege to Russell, 5/24/1903; *WB*, 165; Frege to Marty, 8/29/1882; *PW* 232/*NS* 250–1; 'Logik in der Mathematik' (1914)). These repeatedly and explicitly make the point that the denotations of terms are of interest to us only because of our interest in the denotations, in fact the truth-values, of sentences. Indeed, the remarks constitute a fair approximation to the slogan of *Foundations* that only in the context of a sentence do words have a *Bedeutung*.

Further, the implication of the same passages is that our interest and confidence in the truth of sentences that contain terms justifies our interest and confidence in the terms' having the denotations that they are commonly taken to have. This implication appears to echo and perhaps even sharpen the motivation for one of the uses to which Frege put his contextualism in *Foundations* (sections 60, 62)—defending ontological commitment to objects (numbers) in the absence of an intuitive, imagistic, or causal relation to them. Only the general principle underlying this use is suggested in the argument for Theses (a) and (b). But it is clearly indicated: ontological commitment to the denotation of terms is justified insofar as we are justified in acknowledging the truth of sentences that contain them. It is noteworthy that these developments of Frege's contextualist thinking occur in arguments for Theses (a) and (b), which are in turn embedded in arguments for Theses (c) and (d)—the very theses that Dummett holds prevented Frege from maintaining the prominence of sentences in his contextualist principles.

What then are we to say of the considerations Dummett draws from *The Basic Laws of Arithmetic* to support his view that Theses (c) and (d) undermined Frege's commitment to the centrality of sentences in logical theory? Let us begin with the passage Dummett cites from section 10 of *Basic Laws* that states a weakening of the context principle, one that gives no special prominence to sentences. What Frege writes is as follows:

> With this we have determined the courses of values so far as is here possible. As soon as there is a further question of introducing a function that is not completely reducible to already familiar functions, we can lay down what value it is to have for courses of values as arguments; and this can then be regarded as much as a determination of the courses of values as of that function.

A similar passage occurs in section 29:

> A proper name has a denotation if the proper name that results from that name's filling the argument places of a denoting name of a first-level function with one argument always has a denotation, and if the name of a first-level function of one argument that

results from the relevant proper name's filling the ξ-argument-places of a denoting name of a first-level function with two arguments always has a denotation, and if the same holds also for the ζ-argument-places.

These remarks do indeed state a kind of context principle for fixing term denotation—one that does not give prominence to sentences. First-level concepts are not singled out from among the first-level functions. (Part of the reason for this derives from a particular problem that Frege raises in section 10 about the interpretation of his Axiom V. I shall not go into this point here since it would require substantial stage-setting.)

Although the principles just quoted do not give prominence to predication over functional application, or to sentences over terms, they are unquestionably compatible with the view that ultimately it is the use of a subclass of 'terms', the sentences, that counts in justifying interest in term, denotation and confidence in identifying the denotations of course-of-value terms. Dummett is right to note that Frege does not explicitly draw this distinction in *Basic Laws*. He is probably also right in holding that Frege's not doing so is partly explained by his commitment to counting sentences as falling in the same syntactical category as terms. But it does not follow that Frege had lost sight of the philosophical motivations underlying the formal system that he repeatedly stated in other writings during the same period. The circumstance bespeaks a lack of perspicuousness in the formal system—the price of the various economies Frege prized. But it does not evince a major philosophical turn away from the centrality of sentences, ultimately judgment, in motivating logical theory.

I think that the main reason Frege gives no special prominence to sentences over terms in sections 10 and 29 is that to make intelligible the primacy of concepts (or predication) in fixing term denotation, he would have had to have entered on an excursus into his philosophy of language. Such an excursus would have been incongruous in the context of the book as whole, where philosophical discussion was held to a minimum. The strategy of *Basic Laws* is ruthlessly to suppress discussion of philosophical ideas and motivations, except where they are essential to understanding the formal system and the proofs themselves, or where they bear directly on mathematical practice (as in the case of the discussions of definition and consistency). Where philosophical ideas intrude, they are presented tersely and in summary fashion. Except for the polemical introduction, the book is steadfastly mathematical.

The chief consideration that Dummett relies upon for holding that (c) and (d) undermined Frege's commitment to the centrality of sentences in logical theory is that 'the whole thrust of [Frege's] logical doctrines' was 'to recognize no difference in the kind of logical powers that different expressions have save as were explicable by a difference in logical type' (*IFP* 371–2). Since by Thesis (c), sentences and terms are of the same logical type, it follows that they can have no difference in logical power. Dummett admits that Frege never states such a principle. But he holds, ‘it is implicit in his whole procedure; nothing could illustrate it more aptly than the fact that, in the logical system of *Grundgesetze*, no distinction exists between sentential and individual variables . . . ’ (*IFP* 372).

The evidence of the numerous passages that we cited six paragraphs back indicates that this principle must be severely qualified. Although sentences and terms are of the same logical type, according to Frege, some properties in which they differ are of direct and primary importance to logic. Sentences can make assertions and express judgments; terms cannot. The semantical properties of terms are of interest to us only because of our interest in the semantical properties and use of sentences and thoughts. There is no reason for thinking that Frege wanted to deny or suppress these points in his philosophical writings. As we have seen, they are prominent in his post-*Foundations* work.

In fact, sentences and terms are not everywhere interchangeable even within the formal system Frege presents in *Basic Laws*. So in a further sense, they do not have the same ‘logical powers’ despite the fact that they are of the same ‘logical type’. Only sentences can follow the vertical judgment stroke in Frege’s syntax; ordinary terms cannot. This important point requires detailed explication. I shall develop it by reference to a further consideration that Dummett adduces in favor of his view.

Dummett notes that whereas in the *Begriffsschrift* there is a restriction in the formation rules against placing the horizontal or content stroke before anything other than an expression with judgeable content—anything other than a sentence—in *Basic Laws* this restriction is relaxed. In the latter book, the horizontal may occur before any term (or sentence) yielding ‘a name of a truth-value, of the True if the original expression named the True and of the False in all other cases’ (*IFP* 371). Dummett does not explain what he takes the significance of this fact to be. But it may suggest to the unwary that Frege’s system was set up so as to allow one to ‘judge’ (impossibly) the contents of terms. For example, both ‘−5’ and ‘⊦┬5’ are grammatical expressions in Frege’s logic, where the shorter vertical line in the latter expression represents negation. (Cf. the end of *Basic Laws*, section 6.)

This reasoning would be quite mistaken. (I do not claim that Dummett employs it.) In fact, Frege’s use of the horizontal in *Basic Laws* constitutes one of the subtleties of the book that suggest that Frege was keeping his philosophical motivations in mind. I do not see that the use supports Dummett’s view in any way. To begin with, although the horizontal may apply to any name, it is itself a concept expression: a function from objects to truth-values, as Frege explains (*BL*, section 6). Informally, the horizontal means ‘is the True’. Concept expressions are predicates, and concepts are the denotations of sentential parts (e.g. *PW* 119, 193/*NS* 129; ‘Ausführungen über Sinn und Bedeutung’ (1892–1895); *NS*, 210; ‘Einleitung in die Logik’ (1906)).

Thus the expression ‘– 5’ is a *sentence*, though a false one. It says that 5 is the True. ‘⊢┬ 5’ represents an assertion that 5 is not the true.

Now the vertical judgment stroke can be applied only to the horizontal, content stroke. So it is built into Frege’s system, however discretely, that only sentences, not ordinary terms, may be asserted. Only the senses of sentences, thoughts, may be marked as judged. Frege himself makes the point:

> I distinguish the *judgment* from the *thought* in this way: by a judgment I understand the acknowledgment of the truth of a *thought*. The presentation in the concept script (*begriffsschriftliche Darstellung*) of a judgment by use of the sign “⊢”, I call a statement (*Satz*) of the concept script, or briefly a statement. . . . Of the two signs of which “⊢” is composed, only the judgment stroke contains the act of assertion. (*BL*, section 5)

Judgments acknowledge the truth of a thought, and thoughts are said, over and over again throughout the period and afterward, to be characteristically expressed by declarative sentences: ‘The proper means of expression of a thought is a sentence’ (1897) (*PW* 126, 131/*NS* 137, 142–3; ‘Logik’ (1897)). (Cf. also *PW* 129, 138, 167, 174, 197–8, 206, 216, 243/*NS* 140, 150; ‘Logik’ (1897); *NS* 182; ‘Über Euklidische Geometrie’ (1899–1906?); *NS* 189; ‘17 Kernsätze zur Logik’ (1906 oder früher); *NS*, 213–14; ‘Kurze Übersicht meiner logischen Lehren’ (1906); *NS*, 222–3, 234, 262; ‘Logik in der Mathematik’ (1914); Kl 511, ‘Der Gedanke’/*KS* 345; O 61; G & B ‘S & R’ 64/*KS* 150; O 34–35; etc.)[10]

The result of attaching the judgment stroke to a sentential expression, begun by the horizontal, asserts something, but it is not a term: ‘The judgment stroke cannot be used to construct a functional expression; for it does not serve, in conjunction with other signs, to designate an object: “$\vdash 2 + 3 = 5$” does not designate anything; it asserts something’ (G & B, ‘F & C’ 34/*KS* 137; O 22 n). The vertical judgment stroke is not a function sign, but is the sign of an act—judgment or assertion—an act that applies only to thoughts or sentences. (This is why one cannot substitute a singular term denoting truth for the sentence beginning with the horizontal in the expression ‘$\vdash 2 + 3 = 5$’ (which would yield the ungrammatical ‘| the True’).) It is here that the distinction between sentences and terms finds its representation within *Basic Laws*.[11]

[10] It is true that Frege writes in *Basic Laws*, section 2, ‘The sense of a name of a truth-value I call a thought’. This might seem to be intended to include all terms (not just sentences and their nominalizations), as long as the term denotes a truth-value. I think that there is no independent evidence that this was Frege’s intention. The remark is illustrated only by sentences. As we have just seen, three sections later, Frege says that thoughts are what are judged. And his system allows the act of assertion or judgment to apply only to sentences or what they express. Thus I think that the remark in section 2 is a slip encouraged, to be sure, by the formal assimilation of sentences to terms.

[11] Furth errs in calling ‘⊢’ (rather than the vertical alone) the judgment stroke, a distinction critical to the points we have been making. But he gives an excellent account of the role of the notion of assertion in *Basic Laws* (cf. *BL*, pp. xlviii–lii).

The change regarding the grammar of the horizontal that Frege makes between *Begriffsschrift* and *Basic Laws* is partly motivated by the grammatical assimilation of sentences to terms. But this motivation is less important than one might think. For in one sense the grammatical assimilation of sentences to terms was *already present* in *Begriffsschrift*. Insofar as this is so, the view that adoption of the position effected a major change in Frege's later philosophy of language is rendered further implausible. In *Begriffsschrift*, section 3, Frege writes:

> A language is imaginable in which the sentence 'Archimedes perished at the capture of Syracuse' would be expressed in the following way: 'the violent death of Archimedes at the capture of Syracuse is a fact'. Here one can, if one wishes, distinguish subject and predicate; but the subject contains the whole content, and the only purpose of the predicate is to present this as a judgment. Such a language would have only a single predicate for all judgments, namely 'is a fact'.... Such a language is our *Begriffsschrift*, and the sign '⊢' is its common predicate for all judgments.

Here Frege is primarily intending to make the point that the subject–predicate distinction of natural language has no comparable importance in his logical theory. But the passage also indicates a more radical point of view that the 'content' of the first sentence can be completely captured by the subject, a *term*, in the second. The sign '⊢' is seen in *Begriffsschrift* as a predicate that adds nothing to the content of the term to which it applies. This viewpoint contains more than the germ of Frege's later commitments to the grammatical assimilation of sentences to terms and to the redundancy conception of truth.

The changes from this position in *Begriffsschrift* to his later stance in *Basic Laws* are fairly easy to separate out. In the first place, Frege more clearly distinguished in the sign '⊢' an element corresponding to judgmental force and an element corresponding to the expression 'is a fact' or 'is a truth'. (The running together of force with semantical attribution occurs elsewhere in the *Begriffsschrift*. Cf. for example the semantics given in section 5.) Thus, the vertical judgment stroke represents judgmental force, and the horizontal alone comes to represent a semantical predicate, such as 'is a fact' or 'is true'. Presumably this distinction is accompanied by the rejection in 'Function and Concept' of the *Begriffsschrift* view that the sign '⊢' is a predicate (G & B, 'F & C' 34/*KS* 137; O 122 n—quoted above). On the other hand, the horizontal, taken alone, is a predicate whose meaning is similar to that of 'is a fact'.

Distinguishing force from predication in the sign '⊢' probably made it easier for Frege to relax the *Begriffsschrift* restriction against following the sign '⊢' with anything but a judgeable content. What was asserted need not be just what followed the horizontal, it could be the predication of the horizontal onto what followed it. As I have noted earlier, Frege had already in *Begriffsschrift* come to view predicates as function signs (section 9). Given that he was also already treating the grammar and 'content' of sentences as equivalent to that of terms that nominalize those sentences, it may have seemed a small step to allow the horizontal to be functionally applicable to all terms, simple and complex, clausal and nonclausal. The vertical judgment stroke could still only apply to sentences, the expression of something judgeable.

What seems to be the large step in this development, in addition to the distinction between force and predication, is the semantical clarification that Frege achieved. The semantical standpoint developed in Theses (c)–(d) in effect answered the question of what the arguments and values of the horizontal should be. The redundancy conception of truth is the natural offspring of this semantical development and the viewpoint expressed in *Begriffsschrift* section 3. The universal predicate 'is a fact' gives way to the universal predicate 'is the true'. But in neither case does the predicate add to the 'content' (sense) of what follows.[12] The predicate does not change the sense (or denotation) of the results of ordinary predication.

Thus the grammatical change in the restrictions on the horizontal between the *Begriffsschrift* and *Basic Laws* is not really a change from allowing only sentences to occur to allowing terms to occur. It is from allowing only the occurrence of terms that nominalize sentences to allowing all terms. Thesis (c) played a role in motivating this change. But the developments associated with the changed use of '⊢' that seem most significant are different. The significant developments are Frege's drawing the semantical consequences of viewing predication as functional application (not the mere viewing of predication as functional application, which is already present in *Begriffsschrift*), and the clear distinction between judgmental force and predication. This latter development, and the prominence Frege gave to truth and judgment in motivating logical theory, undermine any claim that the grammatical assimilation of sentences to terms deprived him of his insight into the basic role of sentences in logical theory.

III

The claim that truth-values are objects inevitably suggested to Frege the question 'Which objects?' A parallel question had arisen in *The Foundations of Arithmetic*, once it had been concluded that numbers were objects. Frege was sensitive to the initial possibility that the answer to the latter question might be no other than 'why, the numbers—0, 1, 2 . . . '. Similarly, the truth-values might turn out to be specifiable only as truth and falsity. But Frege's belief in logicism drove him to seek a different answer in the case of numbers. Similar forces were at work in his views on truth-values.

[12] With one exception. When in *Basic Laws* the horizontal applies to expressions like '5' or 'the course of values such that . . . ', which are not sentences, the sense of the result of the application is different from the sense of the argument expression. The former is a thought; the latter is not. There is no analog in *Begriffsschrift* since the horizontal only applied to judgeable contents.

In *The Basic Laws of Arithmetic* the truth-values are identified with particular logical objects, particular extensions of concepts. The reasoning behind this identification is the subject of this final section. Our discussion of this subject must be more conjectural than that of Theses (a)–(d) because Frege wrote very little directly about it. Nevertheless, by piecing together different strands of his views, it is possible, I think, to weave a pattern that has some interest, and even a kind of blemished beauty.

One reason why Frege's reasoning is interesting is that it sheds light on his conception of logical objects. Another is that it is critical to assessing what sort of realism Frege maintained with regard to such objects, and with regard to numbers. Each of these issues is quite difficult and complicated. I shall begin with some very rudimentary background for the realism issue.

A common and straightforward story about Frege's realism goes as follows:

> Frege believed that the numbers are genuine, existing abstract objects. He thought, however, that number theory is reducible to logic. He proceeded to try to show this by constructing a logic containing a version of set theory. He gave definitions, within the logic, of the primitive expressions of number theory, and tried to derive the axioms and theorems of number theory within his logic. Since he had a realist attitude toward the ontologies of the languages of both number theory and logic, and since he regarded numbers as particular objects, he thought that there was but one way to construct the definitions of numerical expressions within his version of set theory. As it turned out, Frege's set theory is inconsistent; and for any viable set theory there are an infinite number of ways of defining arithmetic within it. So even if his set theory had been consistent and even granting that set theory is logic, Frege's logicism and his realism about the numbers are, if not incompatible, at least deeply at odds.

There is much that is right about this familiar recitation. But it seems to me misleading in some fundamental ways. The first derives from Frege's attitude toward all language other than his own concept script. It is well known that Frege thought that natural language is defective for the purpose of expressing thought. But he also thought the same of mathematics itself. Within mathematics, the problem was partly just that the language had not been given logical form. But vagueness was also a problem. Frege repeatedly notes that the content or sense of the term 'number' is not adequately or sharply grasped by even the most competent mathematicians. Other fundamental, long-standing arithmetical terms are afflicted by vague usage. In the first section of *The Foundations of Arithmetic*, he writes:

> The concepts of function, continuity, limit, and infinity have been shown to stand in need of sharper determination. Negative and irrational numbers, which had long since been admitted into science, have had to undergo closer scrutiny of their credentials. In all directions these same ideals can be seen at work—rigor of proof, precise delimitation of extent of validity, and as a means to this, sharp grasp of concepts.[13]

[13] The words 'determination' (for '*Bestimmung*') and 'grasp' (for '*zu fassen*') in this translation replace two occurrences of the word 'definition' in J. L. Austin's otherwise good translation of this

In the introduction to the book, referring indirectly to the concept of number, Frege writes, 'Often it is only after immense intellectual effort, which can continue over centuries, that a concept is successfully recognized in its purity, stripped of foreign coverings that hid it from the eye of the intellect' (p. vii). (Cf. *KS* 122; 'Über das Trägheitsgesetz', O 157–158.) Clothing, covering, veiling are standard Fregean metaphors for the interferences of language in thought. This theme runs throughout the book. Indeed, the book may be fruitfully read as an attempt to remedy the inadequacies of language (primarily mathematical language) for ideally rational conceptualization and thought. If one substitutes 'perception' for 'language', one has the schema for the traditional rationalist program for freeing the intellect from non-rational factors.

Any number of senses and denotations are compatible with the conventional significance of vague terms. It is clear that Frege thought that conventional mathematical usage left mathematical terms vague. That is, what a conventionally competent speaker masters does not fix a definite sense or denotation. Frege thought that his logicist program was required to uncover the senses and denotations of number words. So strictly speaking, defining arithmetical terminology is not a matter of capturing linguistic meaning as we commonly understand it, but of uncovering supplementations of such meaning so that the terminology can be seen to have a definite sense and denotation. Sense lay beyond or beneath conventional significance. (These points are discussed and substantiated in some detail in Burge 'Frege on Extensions of Concepts' (Chapter 7 below).)

It would be a mistake to infer from this point that Frege held that defining numerical terminology involved stipulating meanings for it. The definitions had to respect mathematical practice. Moreover, I think it plausible that Frege thought that only one set of definitions (his) respected all relevant philosophical considerations. Frege's point is that by merely understanding the linguistic meaning of ordinary mathematical language, by being a competent participant in the conventions governing it, one did not *thereby* attain a completely adequate grasp of numerical concepts; one did not *thereby* secure completely definite denotations for number words.

Thus if Frege did think that there was a single, correct set of logicist definitions, it was not because the ordinary conventional meaning of mathematical language and standard mathematical practice allowed only one such set. Since mathematical usage was vague, it admitted of many sharpenings. Rather the unique propriety of a set of logicist definitions would have to depend partly on philosophical considerations, considerations attendant on the logicist program.

Thus Frege's realism about the ontology of ordinary mathematics is more subtle and qualified than the familiar narrative spun above suggests. His

passage from *FA*. Austin's choices may obscure the fact that Frege thought not only that we need a reduction of arithmetical terminology to other terms, but that we need a better grasp of the notions that such terminology expressed.

realist attitude toward the language of mathematics is tied to and supplemented by the assumption that his logical theory gives a proper account of the objects and functions in the mathematical world.

Even with respect to his logical language, his realism is not completely unqualified. Not every logical constituent of the language corresponds to an item in reality. For example, the function sign negation does not in general correspond to a thought component (*PW* 149–50, 185, 198/*NS* 161–2; 'Logik' (1897); *NS*, 201; 'Einleitung in die Logik' (1906); *NS*, 214; 'Kurze Übersicht meiner logischen Lehren' (1906); cf., however, G & B 'Negation'; 131–2/*KS* 374–5; O 155–156).

Nevertheless, Frege thought of his logic as a tool for discovering the nature of the mind-independent world, at least that portion of the world with which mathematics is concerned: ' . . . The mathematician cannot create something at will, any more than a geographer can; he too can only discover what is there and name it' (*FA* 107–8). The theme runs through *Foundations*, his correspondence with Hilbert, and his attack on the formalists; it emerges in the introduction to *Basic Laws*, and it recurs in his late writings. Functions, thoughts, and (at least until the despair over Russell's paradox) courses of values are among the charter members of the mind-independent world. Although this traditional 'realist' interpretation has been questioned now and again, I think it fundamentally secure and will not argue for it in general terms here.[14]

A second way in which the familiar account of Frege's realism that I recited above is misleading concerns the references to set theory. I will not go into this complicated and somewhat obscure matter here. (I have discussed it in Burge, 'Frege on Extensions of Concepts' (Chapter 7 below)). I will just say bluntly that it is a mistake to think of Frege's theories of courses of values and extensions of concepts primarily in terms of what we now know as set theory. This is not because Frege's theory turned out inconsistent. It is because he consciously and repeatedly argued against the basic intuitions that underlie the iterative conception of sets, and because the fundamental intuitions underlying his own theory have only scattered echoes in mainstream set theory and even in the various nonstandard versions. The question of whether and in what sense Frege's logicism and his realism about numbers are affected by the multiplicity of models of arithmetic within set theory is clearly bound up with these matters. This is a question that I shall not attempt to settle here.

If Frege's extensions of concepts were thought of set-theoretically, the axiom of foundation would be everywhere violated. This is a prime reason for feeling queasy about the set-theoretic explication. This reason is a corollary of a more fundamental one—the primacy for Frege of concepts over classes, or

[14] For a discussion that rebuts recent attacks on this interpretation, and with which I am in broad agreement, see Michael Dummett, *IFP*.

courses of values. It is better to think of extensions of concepts visually in terms of a geometrically represented graph, or yet better as the total (abstract) event of matching each of the arguments with their truth-values, one by one. Needless to say, these heuristics give one only a vague start at the notion. (As logical objects, extensions of concepts were not supposed to be dependent for their conception on intuition or vision.) Frege never achieved a clear conception of extensions of concepts that accorded with his philosophical and mathematical preconceptions. It is arguable that no one else has either. So we must be willing sometimes to grope along in the dim afterglow of Russell's devastating paradox if we are to follow the course of his reasoning.

The common thrust of the two main caveats that we have entered in the familiar story about Frege's realism is that we need to be sensitive to the role of his philosophical considerations, beyond what we now think of as standard mathematical (arithmetical and set-theoretic) practice, in assessing the character and plausibility of Frege's realism about mathematical and logical objects.

How do truth-values figure in all of this? They are fundamental in Frege's notion of the extensions of concepts, a subset of which constitutes the primary logical objects. Since concepts are functions from arguments to truth-values, and extensions of concepts are courses of such values, the truth-values chart the courses. Since the numbers are certain extensions of concepts and since truth-values thus figure essentially in the ontology of the numbers, consideration of them is inseparable from consideration of Frege's realism about numbers.

But there is a more specific reason for ontological interest in truth-values. Frege identifies not only the numbers but the truth-values themselves with courses of values, extensions of concepts. Frege indicates that his identification of the truth-values with the specific courses of values he chooses in section 10 of *Basic Laws* is *arbitrary* relative to the axioms of his logical theory. Any other choice would have been equally consistent with those axioms. Getting straight what Frege means in this passage, which we shall scrutinize shortly, is critical for understanding his whole philosophical standpoint.

On its face, the passage in section 10 of *Basic Laws* suggests that Frege's theory of truth has a large stipulative component. It also suggests that Frege's ontology of the numbers contains a massive dose of arbitrariness, and that he was aware of this. If these suggestions are correct, then the traditional view of Frege's realism and of the intentions governing his logicist project must suffer substantial qualification. For different choices as to how to identify the truth-values with extensions or concepts would yield different accounts of which objects the numbers are.[15]

[15] This interpretation has been urged in an article, containing many interesting secondary points, by Paul Benacerraf, 'Frege: The Last Logicist', *Midwest Studies*, 6 (1981). Benacerraf cites not only

I believe that these initially plausible suggestions are mistaken. Although I shall stop short of a general discussion of how Frege regarded his definitions of the numbers, I shall argue that his identifications of the truth-values were, and were known to be, not in the least arbitrary, but supported by reasons. To understand these reasons, we must consider the philosophical context in which Frege conceived his logicist program. The reasons are not narrowly mathematical. Their failure to appear in *Basic Laws* accords with the predominantly mathematical emphasis of the book.

In section 10 of *Basic Laws* Frege correctly argues, first, that whether or not one or both truth-values are courses of values and, second, which courses of values they are, granted that they are courses of values, is left completely undecided by the axioms of his logical system (in particular by Axiom V). He concludes:

> Thus without contradicting our setting ‘$\acute{\varepsilon}\Phi(\varepsilon) = \acute{\varepsilon}\Psi(\varepsilon)$’ equal [in denotation] to ‘$(x)(\Phi(x) = \Psi(x))$,’ it is always possible to stipulate that an arbitrary course of values is to be the True and an arbitrary different one, the False. Accordingly, let us lay down that $\acute{\varepsilon}(-\varepsilon)$ is to be the True and that $\acute{\varepsilon}(\varepsilon = \sim (x)(x = x))$ is to be the False.

‘$\acute{\varepsilon}\Phi(\varepsilon) = \acute{\varepsilon}\Psi(\varepsilon)$’ is read ‘the course of values of the concept Φ is identical with the course of values of the concept Ψ’; ‘$\acute{\varepsilon}(-\varepsilon)$’ is read ‘the course of values of the concept *is the True*’; ‘$\acute{\varepsilon}(\varepsilon = \sim (x)(x = x))$’ is read ‘the course of values of the concept *being identical with the truth-value of not all objects’ being self-identical*’. This passage certainly appears to support the view that Frege’s choice is a matter of stipulation.

And, of course, it is in two ways. The choice involves ‘stipulation’ relative to the commitments of natural language. As Frege often notes, ordinary uses of ‘true’ and ‘false’ do not explicitly commit themselves to truth-values as objects at all. The commitment is promoted by logical theory. Thus relative to natural language use, the identification of any object as one of the truth-values is arbitrary. Frege’s choice also involves stipulation relative to considerations of consistency with Axiom V, and the other axioms of his system. Within a context in which mathematical consistency is all that matters, the choice is arbitrary. But let us broaden the context.

On numerous occasions outside of *Basic Laws* Frege holds that consistency does not suffice for truth. Frege repeatedly defends this view in opposition to early expressions of the model-theoretic viewpoint toward mathematics. Frege’s best-known defenses of the view occur in his correspondence with Hilbert (1895–1900) and in ‘On the Foundations of Geometry, I’ (1903) (e.g. *PMC* 48/*WB* 75; Frege to Hilbert, 6/1/1900; *FG* 25–37/*KS* 264–72; O 321–324, 368–375). But the view is already quite explicit in ‘On Formal

section 10 of *Basic Laws* but some passages in *Foundations*. I have discussed these latter in Burge, ‘Frege on Extensions of Concepts’ (Ch. 7 below). I believe that I have shown them not to support the view. Here I shall concentrate on *Basic Laws*, section 10.

Theories of Arithmetic' (1885) (*KS* 110; O 103). And it clearly guides the criticism of formalism in the closing pages of *Foundations* (pp. 104–119).

These latter passages are particularly relevant to our theme.[16] Frege notes that the denotation of 'the square root of -1' is not fixed by mathematical usage prior to the advent in mathematics of complex numbers (pp. 106–7, 110n). He then ridicules the view that one can simply introduce a denotation for the term by mere stipulation or definition (pp. 107–8). One reason Frege gives is that even granted that the purported definition is consistent, one is not thereby guaranteed that there exists an object that satisfies the concepts used in the definition (pp. 108 ff.): in effect, consistency does not entail truth. A second reason is that even if one succeeds in attaching the term to an object and even if one stipulates meanings for the usual mathematical function signs in application to this object that are compatible with those mathematical principles that had been established prior to the introduction of complex numbers, there might be philosophical considerations that militate against the definition.

Let us elaborate the second reason in more detail since it bears directly on the treatment of truth-values in *Basic Laws*. Frege notes that simultaneously with the introduction of new numbers, the meanings of functional words like 'sum' and 'product' are extended. Suppose we choose some object as the denotation of 'the square root of -1', say, the Moon. So the Moon multiplied by itself is -1: 'This explication seems to be permitted because the [denotation] of such a product does not at all arise from the erstwhile denotation [*Bedeutung*] of multiplication, and therefore in extending this erstwhile denotation it [the denotation of the product] can be arbitrarily determined' (p. 110).[17]

Frege goes on to consider multiplication and addition as applied to imaginary numbers, and in so doing capriciously takes the time interval of one second, instead of the Moon, as the denotation of 'the square root of -1'. He summarizes by saying, 'one is tempted to conclude: Thus it is quite immaterial whether i denotes a second or a millimeter or anything else, provided only that our laws of addition and multiplication hold good; that alone is what matters; the rest need not trouble us' (p. 111).

Frege does not accept this position. One point he makes against it, less interesting for our proposes, is that a contradiction may lurk between the definitions and the rest of mathematical theory. There is no evidence that Frege thought that the relevant definitions in fact lead to contradiction. Frege's other objection is philosophical. By letting the interval of a second be the denotation of 'the square root of -1',

[16] I am indebted to Mary Dant for bringing home to me the kinship of these sections of *Foundations* to the writings on Hilbert.

[17] The bracketed substitution of 'denotation' ('*Bedeutung*') for 'sense ('*Sinn*') is specifically suggested by Frege, in the light of his subsequent sense–denotation distinction, in a letter to Husserl of 24 May 1891.

we are bringing something quite foreign, time, into arithmetic. The second stands in absolutely no intrinsic relation to the real numbers. Propositions proved with the aid of complex numbers would be a posteriori judgments, or at least synthetic, unless we could find some other proof for them, or some other [denotation] for *i*. We must at any rate first make the attempt to show that all propositions of arithmetic are analytic. (p. 112)

For two sections Frege develops the theme of not importing anything foreign into arithmetic. And he ends the book by recapitulating it: by explicitly defining the numbers as extensions of concepts, one can avoid importing physical objects or geometrical intuitions into arithmetic (p. 119).

These sections of *Foundations* provide an initial clue to understanding Frege's remarks about truth-values in section 10 of *Basic Laws*. Like complex numbers, truth-values are seen by Frege as introduced (recognized) for theoretical reasons. Their introduction also extends previous mathematical and natural-language usage. And a variety of ontological choices are compatible with that usage. In showing that Axiom V does not fix the denotation of the course of value notation (or of the expressions 'the True', 'the False'), Frege is indicating, as he does in his attack on Hilbert and the formalists, that the (partially interpreted) axioms of a theory do not by themselves fix the objects of the theory (or the senses of its terms). One must have an understanding of the senses and denotations of the terms used in the axioms that is not reducible to mere commitment to the truth of the axioms taken as linguistic objects. (Cf. note 2 and the accompanying text in Section I.) In stating that the stipulations as to the identity of the truth-values are arbitrary relative to previous usage and previously stated axioms, however, Frege is not stating that his stipulations are arbitrary, absolutely speaking.

What we need now is to understand the reasons for the particular identifications Frege proposes in section 10 of *Basic Laws*. I shall approach his position by a process of elimination. The *Foundations* passages already make it clear why the truth-values could not be identified with physical, mental, or geometrical objects. The domain of logic is universal, whereas these objects are the topics of special sciences, and thus their natures are explicated in synthetic propositions. (Cf. *Foundations*, section 3.) Moreover, physical and mental objects exist contingently and are known by a posteriori methods. Logic encompasses the necessary and is fundamentally apriori.

Similar considerations seem to rule out identifying truth-values with any course of values associated with a function or concept denoted by a term that is a primitive of, or is definable with primitives of, one of the special sciences. Thus the extension of the concept *is a cat* (*is an image, is a line*) is inappropriate.

Truth-values could not be identified with senses because of the arguments of Thesis (b) that we considered in Section I. A corollary of these arguments is that such identification would be inappropriate because senses are denoted primarily in oblique contexts, whereas a truth-value is denoted in the expression of any thought. Since, by Thesis (d) truth-values are objects, and since Frege thought no function (or concept) is an object, truth-values could not be identified with functions.

These considerations leave Frege either with identifying truth-values with the course of values of some logical function or with not identifying them with any courses of values, taking them rather as primitive, 'independent' logical objects.

We have been ruling out possible identifications by appeal to what is foreign to logic. In order to proceed further, we need to recall what Frege saw as essential to logic. Logic essentially concerns itself with the laws of truth. As we have seen, Frege sharpened this claim by stating that the laws of truth yield those norms governing ideal assertions or judgments (*PW 252*/*NS* 272 'Meine grundlegenden logischen Einsichten' (1915); Kl 507–8/*KS* 343; 'Der Gedanke', O 59–60). Truth is the 'objective' of judgment; the most general laws governing this objective form the subject matter of logic. Thus the true is in a sense the most basic logical object. We shall return to this point and sharpen it.

The conception of truth as the aim of logic informed Frege's view that logic has an internal ordering. Logic is founded on the propositional calculus, or the calculus of truth-values. Frege repeatedly emphasizes that one of his seminal contributions is to begin in logic with the sentence and to derive an analysis of sentential parts and their various semantical functions only in the context of a semantical analysis that already features logical relations among sentences: the functional calculi are built upon the calculus of truth-values. We have cited various passages to this effect in his arguments for the view that truth-values are the denotations of sentences (Theses (a) and (b)). Frege also makes the point in his earliest and latest writings (*PW* 17, 253/*NS* 18–19; 'Booles rechnende Logik und die Begriffsschrift' (1880/1881); *NS*, 273; 'Aufzeichnungen für Ludwig Darmstaedter' (1919)). The sense of an ordering within logic is perhaps most clearly enunciated in a footnote Frege wrote in 1910 to Jourdain's chapter on Frege in a history of mathematical logic:

> To found the 'calculus of judgments' on the 'calculus of concepts' . . . is to reverse the correct order of things; for classes are something derived, and can only be obtained from concepts (in my sense). But concepts are something primitive that cannot be dispensed with in logic. . . . And the calculation with concepts is itself founded on the calculation with truth-values (which is better than saying 'calculus of judgments'). (*PMC* 192n/*WB* 287n)

The truth-values are on the ground floor of logic—in the ontology of the propositional calculus. They are a 'subject matter' for all parts of logic. In a sense to be explicated, truth is even more basic than falsity. The laws of logic were for Frege 'nothing other than an unfolding of the content of the word "true" ' (*PW 3*/*NS* 3; 'Logik' (zwischen 1879 and 1891); cf. Kl 507/*KS* 343;

'Der Gedanke', O 59–60). Now courses of values were supposed by Frege to be logical companions of functions, and functions are denoted in all parts of logic. Prior to discovery of the paradox, each function was thought to be accompanied by its associated course of values. Logical objects such as courses of values could be canonically specified only through denoting the associated functions. Frege repeatedly emphasizes that (denoting) a course of values is derivative from (denoting) a function. A function sign denotes a function, but its use determines an associated course of values. Frege seems to have regarded sentences containing function signs as ontologically committed to their associated courses of values. (See G & B 27, 49–50/*KS* 130; 'Funktion und Begriff', O 11; *KS* 173; 'Über Begriff und Gegenstand,' O 200; the point is also suggested in *BL* 4/*GG* i. 7 and *PW* 123/*NS* 134; 'Ausführungen über Sinn und Bedeutung' (1892–1895).)

Since Frege saw logic as having a fixed order, his taking truth-values as part of the ontology of the propositional calculus meant that his specification within logic of these objects could not depend on concepts whose specification was conceptually derivative.

These considerations suggest an argument for taking truth, or the True, to be specifiable in terms of a concept that is primitive within the propositional calculus, assuming that truth is a course of values. Since all logic is concerned with truth and is in fact the unfolding of the laws of truth, and since truth is an object, truth must be in the ontology of all parts of logic—in particular the most fundamental part, the propositional calculus. For a course of values to be in the ontology of the propositional calculus it is necessary and sufficient that it be specifiable in terms of functions denoted in the propositional calculus. So assuming that truth is a course of values, it must be specifiable in terms of functions within the propositional calculus. The only such functions are concepts.

One might worry about the argument, both as a reconstruction of Frege and as a substantive proposal, by concentrating on the second premise. Let us assume with Frege that to be specifiable at all, a course of values must be specifiable in terms of its associated functions. But why is it necessary, in order to be in the ontology of the propositional calculus, for a course of values to be specifiable in terms of functions denoted by expressions in the propositional calculus? And is denoting certain functions really sufficient for being ontologically committed to their associated courses of values?

Of course, with the hindsight that we have gained from the semantical and set-theoretic paradoxes, both questions can be pressed. And I would not wish to defend a Fregean answer to either. It is arguable that sometimes the ontological commitments of a language (say, those involved in giving a semantical theory for it) are specifiable only in a stronger metalanguage. And since Frege's Axiom V is inconsistent, it is sometimes the case that commitment to a given function is not sufficient for commitment to an associated extension (a course of values). This problem makes it plausible to deny that in using function expressions one has dual commitments, to a function and a course of values, even in cases where there *is* an associated course of values. One can make one's commitments separately. But these problems are bound up with larger problems concerning Frege's conceptions of truth and courses of values—particularly with the inconsistency of his system. They do not undermine the argument we gave as an interpretation of Frege.

How would Frege answer the two questions about the second premise? To the first, I think that he would reply that a language that lacked the concepts (or could not denote the concepts) needed to specify its ontology would be logically deficient. Such a language could certainly not serve to express the fundamental part of logic. A logical language that could not specify its own ontology would be dependent on some other language to specify its primary subject matter, truth. Thus it could not be the fundamental language for expressing the laws of truth. But Frege regarded logic as an ideal language of thought; the fundamental part of that language should be expressively complete for its own purposes.

This point should be seen in the light of Frege's redundancy conception of truth. Frege believed that semantical discourse about a language could add nothing to what could already be said in the language itself. It is obvious that he did not anticipate the sorts of considerations that lead us to be cautious about claims that a language must be able to specify its own ontology.

The second question about the second premise of our argument is less interesting insofar as it bears on the interpretation of Frege. The truth-values are in fact in the ontology of his propositional calculus and they are in fact specifiable in terms of functions (concepts) that are denoted in the propositional calculus. Since the truth-values are courses of values in his view, and since he thought courses of values could be specified only in terms of their associated functions, it is hard to see any ground for denying that he thought that being so specifiable was sufficient for being in the ontology of the language. (For a discussion of texts that suggest that Frege thought that language always carried dual commitments to concepts and their associated extensions, see Burge 'Frege on Extensions of Concepts' (Chapter 7 below), sections III–IV.)

I have argued that Frege conceived truth as the subject matter of the most basic part of logic, and that truth had to be specifiable by means of the primitive predicates in the propositional calculus. Frege's conception of an ordering within logic motivates a corollary restriction. Truth should not be the extension of a concept (or course of values of a function) whose specification is in any way conceptually derivative. Truth must be specifiable in terms of a concept whose own specification need not presuppose other types of concepts. This consideration would seem to rule out identifying truth with any course of values of a second-level function (including second-level concepts). For the second-level functions are introduced in logic only after, and only in terms of, first-level functions. The consideration also seems to rule out identifying truth with any course of values of a function, of any level, whose canonical explication presupposes specification of second-level functions.

The effect of these exclusions is substantial. Two large categories of logical objects are barred as candidates for being identified with the truth-values. First, the courses of values that are not extensions of concepts are excluded. There is only one purely logical, primitive first-level function sign that is not a concept sign, or predicate, in Frege's logic. This is the description operator, and its explication presupposes specification of the course-of-values operator, which is second-level (at least!). Of course, the course of values associated with the course-of-values function sign '$\grave{\varepsilon}$' is excluded. For the function it denotes is (at least) second-level. A second category of logical objects that is excluded consists of the extensions of concepts with which the numbers are identified. For these are extensions of second-level concepts. In fact, the definition of 'equinumerous', which is essential to specification of the numbers, depends on second-order quantification—i.e. third-level concepts.

So, if the truth-values are to be identified with courses of values at all, and if Frege's philosophy of logic is to be respected, it appears that they must be identified with the extensions of logical concepts that are first-level. Such concepts ought to be denoted in the propositional calculus.

What underlies the path of exclusion that we have so far followed, and indeed what guides us to our destination, is Frege's conception of truth. The argument we have just given is rather speculative, considered as an interpretation of Frege's own reasoning. The considerations that follow are much less so.

Frege identifies truth (the True) with $\grave{\varepsilon}(—\varepsilon)$. '—', the horizontal, denotes the concept that maps truth onto truth and everything else onto falsity. Thus the horizontal denotes the concept under which only truth falls. Truth is the extension, or course of values, of this concept. Falsity is the extension of the concept under which only falsity falls.

Now this identification can seem to be a piece of artifice if one thinks of courses of values purely in set-theoretic terms. From this perspective truth (falsity) is identified with its own unit class. (Cf. Dummett, *IFP* 404.) What could be more typical of a mere technical convenience? But the set-theoretic gloss misrepresents Frege's view.

Frege's identification should be seen as the result of drawing out the implications of the two ideas that we have discussed so far, and supplementing them with his redundancy conception of truth. The first idea was that the truth-values are logical objects. Their specification should not 'import anything foreign into logic'. Further, as logical objects, their specification must be derivative from the specification (or denotation) of logical concepts. The second idea is that logic is an ordered unfolding of the laws of truth, where truth is the aim of sentence use within logic. Truth must somehow be the objective and subject matter (object) of all parts of logic, including the most primitive part. In fact, it must somehow be implicated in the aim of every sentence of logic. As we have seen in earlier sections, Frege interprets the logically relevant aim of sentence use in terms of our 'striving after truth'. This aim is revealed in assertion and judgment. So truth must somehow be implicated in the assertive use of every sentence of logic. Putting the two ideas together, we seek a concept in terms of which we can specify truth as a logical object, a concept that is present in the assertive use of every sentence of logic.

It might appear that we have an approximation to the idea that truth must be implicated in the assertive use of every sentence of logic, in Frege's doctrine that every sentence of logic denotes truth or falsity. As we have seen, in Section I, there was a connection in Frege's mind between the point of sentence use and the denotation of sentences. But truth is the aim of logic; falsity, strange to say, is not. This aim is revealed in assertion, not simply in the grammatical form of sentences. The concept in terms of which truth is specified is present in every assertive use of a sentence, whereas the counterpart concept for falsity is not. Thus the specification of truth is philosophically primary. The specification of falsity will present itself as natural once we have understood Frege's specification of truth.

The only concept that fits the requirement of being present in the assertive use of every sentence of Frege's logic is his concept of truth, the concept denoted by the horizontal. In unpublished writing contemporaneous with *Basic Laws*, Frege is quite explicit about the point. He held that what distinguishes 'true' from all other predicates, and what fits it to indicating the aim of logic, is that 'it is asserted when anything at all is asserted' (*PW* 129/*NS* 140; 'Logik' (1897)). In accord with his redundancy conception, Frege held that 'it is true that ——', when filled by a declarative sentence, expresses the same sense as '——' when filled by the same sentence. But this neutrality of sense in predications of 'true' is collateral with the predicate's omnipresence in assertions and judgments. As we have seen in Section II, this idea of the omnipresence of the truth predicate traces all the way back to the *Begriffsschrift* (*B* 3).

The horizontal expresses the notion of truth in Frege's system. It means 'is the True' or 'is the truth' or 'is truth'.[18] It is present in the formulation of every assertion. It may accompany any declarative sentence without adding to its sense. The concept denoted by the horizontal is the only one within Frege's logic that meets the condition set by his redundancy conception of truth. Frege alludes to this condition without fanfare in section 5 of *Basic Laws*, where he notes the equivalence:

$$\Delta = (-\Delta)$$

[18] Almost needless to say, Frege's representation of the truth predicate is not intended exactly to reproduce ordinary language. In particular, '$—\grave{\varepsilon}(—\varepsilon)$' turns out true but has no analog in ordinary uses of 'is true'. Cf. also note 12.

where 'Δ' varies over truth-values. The import of the condition comes clear if one sees sentences as substituting for 'Δ', reads '$=$' (as in such cases one may in Frege's system) as the material biconditional, and reads the horizontal as the truth predicate. The equivalence is the analog within Frege's system of Tarki's truth schema.

Given his redundancy conception, Frege regarded the two sides of the equivalence as having the same *sense* when a sentence is substituted for 'Δ'. (Cf. G & B 63–4/*KS* 149–150; 'Über Sinn und Bedeutung' (1892), O 34.) This suggests that the concept of truth may in a sense be implicated in the ontology of every (declarative eternal) sentence, whether the sentence is asserted or not, and regardless of whether a truth predicate explicitly occurs in the sentence. This conclusion needs the assumption that a sentence is committed to the existence of the extension of a concept *C* if *C* is a denotation determined by (a component of) the sense that the sentence expresses. Since every sentence has the same sense as a sentence in which the concept of truth is denoted, every sentence would by this reasoning be committed to the existence of the extension of this concept—the truth-value truth. Although I think that there is some reason to believe that Frege considered and accepted the relevant assumption, he did not explicitly assert it. (Cf. G & B 27, 49–50/*KS* 130; 'Funktion und Begriff' (1891), O 11; *KS* 173; 'Über Begriff und Gegenstand' (1892), O 200; *BL* 4/*GG* i. 7; *PW* 123/*NS* 134; 'Ausführungen über Sinn und Bedeutung' (1892–1895); the numerous passages where he speaks of the senses of sentences as being decomposable into component senses; and Burge, 'Frege on Extensions and Concepts' (Chapter 7 below).) I shall not defend the attribution of the assumption, however, since I regard it as somewhat speculative.

What is indisputable and important is that Frege forces the truth concept to be *explicitly* denoted by a truth predicate (the horizontal) when a sentence is assertively used in his logic. Thus it is through assertion that the aim and ultimate subject matter of logic are revealed.

The truth concept also expresses Frege's conception of the order within logic. It is the concept in terms of which all others are explicated and understood. In this regard, it is prior to any other concept canonically denoted in the propositional calculus.

In specifying truth as the extension of his truth predicate, Frege is 'deriving' a logical object from a logical concept. Obviously, it is intuitively natural to derive the object truth from the concept of truth (assuming that one wants such an object and that such an object must be derived from a concept). Frege is in effect nominalizing the truth predicate by generalizing on his truth schema (in the light of his construal of the schema in terms of Theses (b) and (d)).[19] But Frege's specification of the object truth is not merely intui-

[19] Frege's method of specifying or "defining" truth (in the extensional, mathematical sense) is a primitive ancestor of Tarski's set-theoretic methods of attaining the same objective. (Frege differs from Tarski in having no ambition to explicate, or provide a vindication for, the concept of truth.) Both

tively natural, granted his assumption that it is a logical object and must be derived from a concept. It also expresses his conception of the point, subject matter and order of logic, and flows from his redundancy conception of truth. It is the only identification that is consonant with his philosophical views.

The evidence we have been considering suggests reconsideration of a pair of long-standing criticisms of Frege's view of truth-values as objects. One is that once truth becomes one object among others, it is difficult to explain what it is about it that makes us want to strive after it, assert it, acknowledge it, and so forth. (Cf. Furth, *BL*, pp. lii–liii. The point is also made by Dummett in various places in *FPL*.) What is so terrific about the relevant object? Frege seems to be inviting us to join a kind of secular religion without explaining the attributes of its god that merit our worship.

We may begin to appreciate the weakness of this criticism by recalling Frege's contextualist arguments for Theses (a) and (b). Denotation of objects with terms is of interest only because of our interest in semantical features of sentences. Interest in sentences derives from interest in norms governing their use in making assertions and expressing judgments. So the practice of using sentences to denote truth-values derives its interest from the role of assertion and judgment in our lives. From Frege's point of view, the idea that we must explain this role in terms of the features that certain objects, the truth-values, have, would be to put the horse behind the cart.

Frege underlines this point by specifying the object truth as what every assertive use of a sentence (and every judgment of a thought) is committed to. We know what the object truth is and what it is like only through reflecting on the fact that assertion and judgment aim at it. To be sure, Frege thinks that the laws of truth, which are laws of 'being' (*Sein*, Kl 507–8/*KS* 342–3; 'Der Gedanke' (1918–19), O 58–59), generate norms for assertion and judgment. (Cf. note 3.) This is because Frege assumes it as obvious that we ought to judge in accordance with logical laws, and because he construes these laws as governing an objective world of objects. But it is not part of his view that we should be able to explain the interest for us of the object truth, or the way that we think about it, independently from consideration of the point of assertion and judgment—to 'strive after truth'.

Frege thinks truth and judgmental force are primitive ideas. And he does not try to explicate or philosophize about the value of 'striving after truth'. But it is foreign to his system, and thus not a pseudo-question concocted by it, to ask what it is about the object truth that engenders our interest in it. The activity of judging and the practice of assertion are primary.

The second long-standing criticism of Frege is closely related to the first. Dummett states it in his thought-provoking article 'Truth':

authors may be seen as summing up or generalizing from their respective versions of the truth schema. Frege's identification of falsity as the extension of the concept *is the false* is the natural counterpart of his specification of truth.

> . . . it is part of the concept of truth that we aim at making true statements; and Frege's theory of truth and falsity as the references of sentences leaves this feature of the concept of truth quite out of account. Frege indeed tried to bring it in afterwards in his theory of assertion—but too late; for the sense of the sentence is not given in advance of our going in for the activity of asserting, since, otherwise there could be people who expressed the same thoughts but went instead for denying them. (*TOE* 2–3)

It is certainly true that Frege said less about the roles of judging and asserting in our lives than one might want in a post-Wittgensteinian climate. But it is not true that Frege leaves our aim at making true statements 'quite out of account' in his exposition of the concept truth. It is this aim, he says in 'On Sense and Denotation', that motivates our asking for the denotations of terms and sentences. Late in life he writes: ' . . . "true" only makes an abortive attempt to indicate the essence of logic, since what logic is really concerned with does not at all lie in the word "true" [since by the redundancy view, it contributes nothing new to the sense of whole sentences in which it occurs as predicate] but in the assertoric force with which the sentence is uttered' (*PW*252/*NS* 272; 'Meine grundlegenden logischen Einsichten' (1915)).

The last main clause of the passage cited from Dummett's criticism is difficult to interpret. But it does not seem relevant to Frege's view. Frege nowhere, to my knowledge, writes or implicates that the sense of a sentence is given 'in advance' of our going in for the activity of assertion. Denotation is motivated and justified by Frege in terms of our 'striving after truth'. Sense is postulated as the way denotations are presented to us in thought. Thoughts, the senses of sentences, are truth conditions; and we are interested in truth conditions because we are interested in arriving at truth. So our attaching senses and denotations to sentences and sentential parts is motivated by their roles in judgment and assertion. Frege's thesis that sentences denote objects is proposed in the context of these motivations, not in contradiction to them.

We should now consider the possibility of taking truth and falsity as 'new' logical objects, not identical with any course of values. Frege would not have introduced a new name for truth, an individual constant, into his system since doing so would have created an inelegant logical connection between his specification of truth and his truth predicate. Moreover, such a move would controvert his view that logical objects are derivative from logical concepts, a view bound up with his contextualist defense of the existence of abstract objects. But suppose that one identified truth with $\backslash\acute{\varepsilon}(-\varepsilon)$ (where the slash is the description operator—cf. section 11, *Basic Laws*)—the unique object that falls under the concept—(*is the true*).

I think that Frege would answer this suggestion by asking rhetorically what purpose a distinction between truth and the extension of the concept *is the true* would serve. He would take $\backslash\acute{\varepsilon}(-\varepsilon)$ to be $\acute{\varepsilon}(-\varepsilon)$. (Compare the way Frege argues against postulating negative thoughts (*PW* 149–50, 185, 198/*NS* 161–2; 'Logik' (1897); *NS* 201; 'Einleitung in die Logik' (1906); *NS* 214; 'Kurze Übersicht meiner logischen Lehren' (1906)).) Prior to the discovery of the paradox, there were no evident logical advantages to the distinction. This point is implicated in note 17 in section 10 of *Basic Laws*.

It is arguable that Frege accepted the intuitive or metaphysical point that a physical, mental, or geometrical object is distinct from any course of values—on the ground that they are obviously distinct categories of objects. (The argument would have to develop out of the central sections of *Foundations*.) But such a point is evidently not applicable to the relation between truth and courses of values. Truth is a necessary, logical object, unlike physical, mental, and geometrical objects; and since it is introduced on predominantly pragmatic grounds, against the grain of prior intuition, there is no force in the claim that we make an intuitive distinction between the object truth and courses of values. From Frege's pre-paradox point of view, I see no purpose to the distinction and so, for him, no point in drawing it.

I believe that the urge to draw such a distinction derives from thinking that has already been informed by the set-theoretic and semantical paradoxes. One wants to insist on a difference in 'level' between basic individuals and higher-order objects. Of course, merely taking truth to be an individual will not suffice to deal with the semantical paradoxes. Frege's redundancy theory of truth and his truth predicate ('is the True') are not capable of representing common uses of the notion of truth, much less of explicating and representing the derivative, indexical, and schematic aspects of the notion, aspects revealed by the paradoxes. (Cf. Burge, 'Semantical Paradox', *The Journal of Philosophy*, 76 (1979), 169–98.) If there is not some pervasive provision for levels of just the sort that Frege ignored, there can be no adequate theory. The notion of truth cannot be adequately represented in terms of a truth predicate that lacks any sort of stratification.

These remarks are, however, anachronistic. They presuppose knowledge that Frege lacked when he wrote *Basic Laws*. Given his redundancy conception of truth, his notion of course of values, his view of truth as an object, and his logicist ambitions, the identification of truth-values that he proposed in *Basic Laws*, section 10, must have seemed uniquely appropriate.

Truth is the basic logical object in Frege's system. It is the object on which the purely logical, first-level functions (including concepts) are initially construed as operating (*BL*, section 31). The logical objects with which the numbers are identified are derivative from second- and third-level concepts that operate on the first-level functions. So the numbers are less fundamental than the truth-values (though all are, of course, seen by Frege as necessarily existent). The reason truth is basic is that it is the object (objective) of assertion and judgment. It is through these activities that the point and the ontology of logic are revealed.

As we have noted, the truth-values formed, at one time, the basis for an argument against scepticism. They also formed the basis for Frege's answer to Kant's dictum that there could be no knowledge of objects without intuitions. Frege held that commitment to mind-independent objects is inseparable from the very act of judging something to be true. The idea is surely the quintessential distillation of the ambitions of rationalist epistemology.

Even apart from these ambitions, the profundity and breadth of Frege's conception must be seen as admirable. What enables Frege's views on truth-value to be of enduring importance is that they are so largely grounded in fruitful and highly articulated insights into the anatomy and function of logic, mathematics, language, and thought. I think it unlikely that we have fully harvested Frege's insights.

Postscript to "Frege on Truth" (2004)

I would like to do four things in commenting retrospectively on 'Frege on Truth' (1986). First and second, I want to correct two mistakes in the paper. Third, I want to discuss critically a very different interpretation of Frege's views on truth and logical consequence, associated with the second correction. Finally, I want to make some remarks about Frege's view of the role of his conceptions of truth in expressing knowledge.

The first correction is easy to state. On page 123 I took the propositional calculus to be the fundamental part of logic. It is true that the propositional calculus is the simplest part. I think, however, that the propositional calculus is not conceptually fundamental.

I believe that Frege would agree. Where he writes, "And the calculation with concepts is itself founded on the calculation with truth-values . . . " (*PMC* 192n/*WB* 287n; notes on Jourdain (1910)), I am inclined to think that he means not that the propositional calculus, as we now know it, can be understood independently of a predicate logic, but rather that our understanding of the behavior of concepts in a predicate logic depends on reflecting on evaluations of the truth-values of sentences and on inferences carried out using sentences.

The propositional calculus is an abstraction from the predicate or functional calculus. I think this because I believe that there are no sentences or propositions that, under analysis, are true or false and that lack predicative form. Sentences or propositions always involve predication. Of course there can be one-word sentences, and the one-word sentences may not *be* predicates. But insofar as they have truth-value, the single words must mask an implicit structure, involving some sort of implicit predicational attribution. I cannot make sense of having truth-value apart from predicational attribution. The propositional calculus treats only logical operations on sentences or propositions. So predicate structure is suppressed or abstracted from. I believe that the arguments of the paper can be revised in relatively obvious ways to take account of this correction.

The second correction has to do with the notions of logical validity and logical consequence. It requires more comment.

On p. 91 I include validity among modal notions. On p. 95 I write of Frege's conception of logical consequence as being in terms of necessary truth. These remarks were mistaken. The mistake was in effect corrected in 'Frege on Apriority' pp. 368–369 where I indicate that Frege understands the validity of logical laws in terms of the generality of their ultimate ground—the ground (such as fundamental logical axioms and inference rules) which justifies them—not in terms of necessity. In fact, Frege understands a truth's *necessity* in terms of the generality of its ultimate ground. He does not treat it as an autonomous notion.

This point is of considerable historical and conceptual importance.[1] Relations among epistemic notions, like apriority, modal notions, like necessity, and formal notions, like generality and structure, have a long history in logical theory. Often the three types of notions were run together. Aristotle's logical theory emphasizes modality over form in understanding logical truth and good deductive inference ('following from'). In the Middle Ages, there emerged an independent tradition in which intuitive conceptions of logical validity and logical consequence centered on form and structure rather than modality. Nearly everyone in this tradition took logical truths and logical consequences to *be* necessary. But the intuitive conceptions were not themselves modal conceptions. The oldest versions of such conceptions explicated logical truth and logical consequence in terms of containment relations among structures in the world that were reflected in logical relations among logical forms of proposition or argument.

These intuitive logical notions eventually received theoretical explication in terms of model-theoretic notions of validity and logical consequence. The intuitive notions, however, are distinct from the model-theoretic notions. They are notions of truth and truth preservation—not notions of truth, or preservation of truth, *in* a model. The intuitive notions are also mathematically inspecific (unlike the model-theoretic notions) and allow for a variety of conceptions of form, structure, and generality.

The key feature of the intuitive notions of logical validity and logical consequence is that they explicate logical truth and deductive truth preservation in terms of logical form and logical structure. They also take logical truths and logical consequence to exhibit a strong form of generality. Usually logical structure is understood to be an aspect of 'the world' or of a subject matter. The generality intuition plays a role in determining what counts as logic, or as a logical constant.

The mistake that I made in 'Frege on Truth' is that I did not recognize the importance of sharply separating this tradition from modal conceptions of logical consequence. It is not that logical consequences are not necessary or that there is anything wrong with modality. It is that the notion of logical consequence can be conceived without appeal to it—in terms of notions of formality, structure, and generality. I believe that this has clearly become the dominant conception in contemporary logic.

Frege is an important representative of the tradition that understands logical truth and good deductive inference in these terms. The near absence of modality in Frege's conception of good deductive inference, and the prominence of these other notions (formality, structure, and generality) is extremely striking and of great historical importance. In fact, I think he made

[1] These historical and conceptual points are developed at much greater length in my 'Logic and Analyticity', *Grazer Philosophische Studien*, 66 (2003), 199–249. There is a discussion there, especially in n. 37, of some of the points about Frege made below.

the key contributions to this tradition that enabled Skolem, Gödel, Tarski, and others to provide the theoretical understanding of the intuitive formal notions of logical truth and logical consequence, and of logic and deductive inference, that flowered in model theory. Frege's non-modal conception of the key logical notions, including his (largely) non-modal explication of rules of inference in terms of assignments of truth-values, his breakthrough work on logical form, his functional explication of the truth of whole sentences (or thoughts) in terms of the denotations of the parts of sentences, and his firm association of logical truth with the generality of the ultimate bases for its justification and the generality (in a different sense) of its application to subject matters, are all fundamental and distinctive components of the intuitive conceptions. In fact, Frege brought these components to a level of formal sophistication that was unprecedented. His elaboration of them formed the basis for subsequent, modern theoretical conceptions of the intuitive notions.

Frege's grip on the pre-theoretical notions of logical truth as logical validity and of good deductive inference as formal logical consequence is deeper than that of anyone prior to those who systematized the intuitive notions in a semantical meta-theory for logic.

Frege's philosophical orientation is quite different from the orientations of most of his successors. Like all others who had the intuitive notions of logical validity and logical consequence who preceded the development of model theory, Frege worked only with the notion of truth. He did not connect that notion with a notion of truth *in* a model (which itself is, strictly speaking, a distinct notion). Whereas domain variation has become a basic tool of model theory, Frege evaluated logical truth entirely by reference to the (actual) world.

This difference is easily overrated. Frege is, in this respect, like every other representative of the intuitive tradition prior to the emergence of model theory. Even most of the early expositions of model theory, including Tarski's original model-theoretic explication of logical consequence, do not employ domains or domain variation.[2]

Frege took the axioms of logic to carry ontological commitment to an infinity of objects (for example, the numbers). He is like Russell in this respect. He is unlike most modern logicians, who take logical axioms to have ontological commitment to the existence of at most one object. This difference in logical conception is compatible with his sharing the non-modal, formal conception of logical truth and logical consequence.

Frege had a different conception of the role of truth and of semantics in logical theory than his successors. He regarded at least certain attributions of truth as being in the object-language and as adding nothing to the *sense* of

[2] Alfred Tarski, 'On the Concept of Logical Consequence' (1936), in *Logic, Semantics, Metamathematics*, (Indianapolis: Hackett, 1983). Cf. Tim Bays, 'Tarski on Models', *The Journal of Symbolic Logic*, 66 (2001), 1701–26.

asserted non-semantical sentences. Frege did not regard the predicative attributions in these occurrences as being executed via a semantical predicate. Rather he understood sentences like '$2 + 2 = 4$' as having the form '2 plus 2's being 4 is the true', where 'is the true' is a simple non-semantical predicate, formalized by the horizontal, that maps the truth-value of '2 plus 2's being 4' onto the truth-value truth (given that the truth-value of the nominalization is truth). It maps the denotations of nominalizations that denote the false onto the truth-value falsehood. And it maps the denotation of every other singular expression onto the truth-value falsehood.

Further, Frege regarded '$2 + 2 = 4$' as having the same sense as 'it is true that $2 + 2 = 4$'. He may have regarded the sense of the nominalization of '$2 + 2 = 4$' as having the same sense as that of the result of applying his predicate 'is the true' to the nominalization. He *clearly* regarded 'it is true that $2 + 2 = 4$' as having the same sense as '$2 + 2 = 4$'.

This redundancy account of the sense of the predicate 'it is true that', which he parsed as 'is the true', is surely false. In fact, the sense and denotation of the predicate 'is the true' (the horizontal), in these uses, is wrapped up with Frege's commitment to extensions of concepts. The predicate denotes a function from truth-values to truth-values—where truth-values are extensions. Since his notion of an extension of a concept is defective, his non-semantical notion of truth, expressed by the horizontal, is doubly defective.

For all that, Frege did have a *semantical* predicate that can be employed so as to fit Tarski's truth schema. This is the predicate 'denotes' (*bedeutet*) applied to the relation between sentences and the truth-value truth. Frege regarded 'denotes' as a two-place relational predicate. Frege frequently uses this predicate in doing some of his deepest and most influential work. Frege's use of this predicate—including his use of it to explicate relations between sub-propositional logical structure and propositional logical structure—grounds my attribution to him of the intuitive notions of logical truth and logical consequence. These are intuitive notions of truth and truth preservation, respectively, explicable in terms of the logical form of sentences and arguments, on one hand, and logical structure to which such sentences are semantically related, on the other.

Let me turn to the third of the three things I want to do in this postscript—discuss an alternative view of Frege's relation to the concept of logical consequence.

Thomas Ricketts holds that Frege 'has no semantic conception of logical consequence in the post-Tarskian sense of "semantic"', and 'lacks any general conception of logical consequence, any overarching conception of logic'.[3] He does not specify what he thinks the post-Tarskian sense of

[3] Cf. e.g. Ricketts, 'Logic and Truth in Frege', *Proceedings of the Aristotelian Society*, suppl. vol. 70 (1996), 124. In other respects, I value Ricketts's sensitivity to respects in which Frege's work is foreign to some elements in the current philosophical climate. But I think that some of the main points of his Frege interpretation are off base and lacking in specific textual support.

'semantic' is.[4] It is obvious that Frege did not present a full-blown model theory.[5] Ricketts seems, however, to be claiming that Frege did not have the *intuitive* notions of formal logical validity or logical consequence, or any other general intuitive conceptions of logic. I want to evaluate some of these claims. I think that they are not well grounded, and thoroughly mistaken.

One reason Ricketts gives for his claims is that Frege never states a 'defining criterion of the logical'.[6] This seems to me not a strong reason. The fact that Frege did not hazard a defining criterion is hardly a ground for holding that he lacked a conception of logic or logical consequence. Frege's use of his two key notions of generality (quantificational generality in the grounds for a truth and generality of application to all subject matters) together with his practice of explaining formal deductive reasoning exhibits a deep conception of logic, pushed through with greater consistency than any predecessor did. These key notions and Frege's use of them do not *constitute* a criterion for logic or even a conception of logical consequence. Frege's theoretical development of formal logic conjoined with his deep insights into the semantical structures associated with logical form are what shows that he has the intuitive formal conception of logical consequence. He did not need to give a definition or criterion.

Ricketts also claims, "Frege is not committed to the claim that every truth couched in topic-universal vocabulary is a logical law."[7] I think that we do not know whether Frege was committed to this claim. He does not explicitly make the claim, to my knowledge. But the claim is not one that every proponent of the intuitive conception of logical consequence would accept. The truth of the claim is not to be taken for granted in any case. For example, in first-order logic $(\exists x)(\exists y) \sim (x = y)$ is not a logical law or even a logical truth, even though it is true and is couched in topic-universal vocabulary. I believe that Frege would have regarded this proposition as a logical truth,

[4] To be maximally specific: I take the intuitive notion of logical validity to be truth grounded and correctly explicable (in any of a number of different ways) in terms of logical form and logical structure. The intuitive notion of logical consequence is preservation of truth grounded or correctly explicable in terms of logical form and logical structure. Logical form is a feature of sentences or propositions. Logical structure is usually taken to consist of objects, properties, relations, functions, sets 'in the world'. The ontology varies with the conception. Grounding or explication can be understood in various ways. In fact, 'logical' can be understood in various ways that are compatible with the intuitive notions. The most common conception of logicality (which Frege represents) involves reliance on the generality of application of the logical constants.

I take 'semantical' to apply to relations between signs, or (or in a broader usage of 'semantical') thought components, and entities that they mean, refer to, express, or the like. Semantics need not be model-theoretic.

[5] It seems to me questionable, however, to claim even that Frege lacked an approximation to the concept of truth in a model. Although he does not give a systematic model theory, and although he does not employ the modern notion of a set (in terms of which models are explained), his reasoning in sections 10 and 31 of *The Basic Laws of Arithmetic* suggests a grip on the theoretical forms of reasoning that came to constitute model theory and the early model-theoretic explications of logical consequence.

[6] Ricketts, 'Logic and Truth in Frege', 124.

[7] Ibid.

but only in virtue of his claim, which he knew to be controversial, that arithmetic is logic and that numbers are logical objects. There are, I believe, other truths (either the continuum hypothesis or its negation) that can be couched in the topic-universal vocabulary of second-order logic but which most logicians do not regard as logical truths. What Frege would have thought about such cases, if he had considered them, is hard to say.

What I want to bring out is that the claim that Ricketts cites, though it might have been attractive to Frege, was actually at least as controversial as particular aspects of Frege's logic. Frege had doubts about these aspects himself.[8] Even if Frege accepted the claim, it would have been bad dialectical strategy for him to assert a criterion for logic or logical truth that was at least as controversial as the most obviously controversial parts of his own logic—the parts on which his logicism about arithmetic hinged.

Much of Ricketts's case for holding that Frege lacked the concept of logical consequence seems to lie in claims that he lacked a truth-conditional semantics—or indeed any semantics relevant to understanding logical notions.[9] Ricketts's claims are associated with the aim of trying to understand Frege's type-restrictions and his conception of truth. But the claim that Frege lacked a truth-conditional semantics is, I think, deeply and obviously mis-

[8] Cf. my 'Frege on Extensions of Concepts, From 1884 to 1903' (Ch. 7 below) for the point about Frege's doubts. Cf. also my 'Logic and Analyticity', n. 23, and Appendix I ('Logic: First-order? Second-order?') of that paper. In the latter paper, I discuss why on natural conceptions of logic and intuitive logical consequence (or logical validity—for sentences or propositions), one should *not* accept the claim that Ricketts takes as criterial for having the intuitive notion of logical consequence. The natural conception of logic that I develop there is not Frege's. The point is that the intuitive notion of logical consequence is not constitutively committed to the claim. The claim would have been attractive to Frege, but whether he is committed to it seems to me unclear. His not explicitly making the claim can be explained by either substantive or dialectical reasons.

[9] Ricketts claims that 'much of what we tend to think of as Frege's semantics is not statable within the framwork of the *Begriffsschrift*. Frege's universalist conception of logic gives it an anomalous status' ('Logic and Truth in Frege', 128). This claim is quite inspecific. Ricketts may just be pointing out that some of Frege's remarks about functions in German or English would be subject to 'the-concept-*horse*' problem. This point would be correct but would have no deep implication for the role of semantics in Frege's understanding of logic. Neither Frege's logic nor his semantics need treat functional expressions as denoting what can be denoted by singular expressions. (I should say that I believe that there is no good reason why functional expressions and singular expressions cannot bear different denotation relations to the same entity, or even the same denotation relation to the same entity, as long as syntactic distinctions are strictly maintained and as long as semantical type-restrictions are placed on what sort of function is being denoted.) Ricketts seems to mean that no semantics that Frege had can be formulated in the framework of the *Begriffsschrift*. Again, this claim would be inspecific. If the point is merely that semantical terms (like '*bedeutet*') are not among the logical constants that Frege employs, that is clearly true. This point certainly could not by itself be used to show that Frege did not have a semantical conception of logical consequence or logical truth. Many who have the intuitive semantical conception, *logical consequence*, would be happy to agree that semantical notions are in a sense not logical constants proper—components of the axioms or even inference rules of logic. If the point is that even taking Frege's semantical notions as non-logical constants, one cannot formulate Frege's semantics in *Begriffsschrift*, this seems to me quite incorrect. Ricketts gives no argument for the claim. I know of no reason to think that the semantics for Frege's first-order functional expressions cannot be formulated in a semantics within the framework of a higher-order logic that distinguishes function-denotation from object-denotation. Function-denotation

guided. Where does such a claim come from? I believe that there are no good textual grounds for this claim. I will discuss the purported bases for it.

One purported ground is a reading of Frege's argument against the definability of truth. Frege's argument goes as follows:

> ... in a definition [of truth] certain characteristics [such as *corresponds to reality*] would have to be specified. And in application to any particular case the question would always arise whether it were *true* that the characteristics were present. So we should be going round in a circle. So it seems likely that the content of the word 'true' is *sui generis* and indefinable.

As it stands, Frege's argument is an enthymeme. Frege does not explain why a definition of truth and meta-level propositional attitudes using the definition must be relied upon in order to decide whether a proposition is true. One might simply make judgments that do not contain the notion of truth and then infer the truth of the judged contents from them. It is unclear to me whether Frege's views are vulnerable to this reply since his views are not explicit.

One might try to save Frege's argument, however, by supplementing it with three further premises. The first is that ordinary non-factive judgments presuppose the notion of truth in some looser way than actually employing or containing the notion. Frege would hold that non-factive judgments at least aim at truth and that this aim constitutes a presupposition of any application of belief or non-factive judgment. I believe that this first premise is certainly true. The second is that a definition gives an epistemic order of priority or justification: the defined term is known to apply through applying the definition. Although this understanding of definitions in the formal sciences is no longer generally held, it was Frege's understanding. Third, if a definition, understood in the way indicated in the second premise, defines a notion that is presupposed in a full account of what it is to apply the definition, the definition is inadmissible. Even if one tried to substitute the definiens for the definition in the presupposition, the definition, understood in the light of

for first-level functions is a two-level relation between an object (a sign) and a first-level function. One simply needs to employ different semantical expressions for the different types of expression being explicated (signs for objects, signs for first-level functions, signs for second-level functions, and so on). In principle, the type theory continues without limit, and the semantics can do so as well.

It is, of course, true that signs are special objects—the province of a special science. But in various places Frege makes it clear that he thinks that the fundamental representations are thought contents and thought-content components. I see no deep reason why Frege could not have regarded semantical predicates—relating thought components and their *determinata*—as logical constants. If he had developed a logic for attributions of propositional attitudes, I think that he may well have done so. (I discuss this point below in the text.) I see no deep reason why he could not have regarded thoughts as logical objects, as long as the thoughts themselves contain no non-logical components.

Of course, no type theory of the sort that Frege proposed has a single notion of generality or universality that applies to 'everything'. Frege clearly envisioned an intensional logic, of the sort that Church developed. (Cf. *Wissenschaftlicher Briefwechsel*, ed. Gabriel, Hermes, Kambartel, Thiel, Veraart (Hamburg-Felix Meiner Verlag, 1976), Frege to Russell, 12/28/1902.) It is far from obvious that he thought that thoughts, in particular those that contain purely logical vocabulary and apply to logical objects and functions, are objects of a special science. They may well have been as universal, for him, as the numbers.

the second premise, would be involved in epistemic circularity. I believe that Frege probably accepted all of these premises. Adding these premises makes the enthymemic argument valid. I believe that this reconstruction is a fairly plausible conjecture about how Frege was reasoning.[10]

Ricketts reads the argument as showing not only that truth is indefinable (its only explicit conclusion) but that truth is not a property. He wants to infer from this result that Frege lacked a truth-conditional semantics. I will later question this latter inference. The inference is the key issue. But I want to begin by discussing Ricketts's reading of Frege's argument as having an additional, unstated conclusion.

I find quite unpersuasive the reading of the argument that takes it to be directed against taking truth to be a property. There is nothing in the text of the argument to support this reading. In Ricketts's exposition the claim depends on the idea that to take truth to be a property is to hold that to recognize a thought to be true must be to recognize a mental representation to have a particular property.[11] But this is just the mistake, alluded to three paragraphs back, of conflating object-level judgments with judgments that a proposition is true. No opponent need accept this view.[12] I see no decisive reason for thinking that Frege made the mistake of attributing such a view to his opponents.

Frege does not draw from the argument against the definability of truth the conclusion that Ricketts attributes to him. Frege never says that the argument shows that truth is not a property. I believe that Ricketts's reading is not only unsupported by the text. It is uncharitable to Frege. The only charitable reconstruction of it that I know of (given above) does not support such a conclusion. Frege did not give the argument against the definability of truth to show that truth is not a property.

[10] For Frege's argument against the definability of truth, see 'The Thought', in *Collected Works* (Oxford: Blackwell, 1984), 353; *Kleine Schriften* 344; O 60; and 'Logik' (1897), in *Nachgelassene Schriften*, ed. Gabriel, Hermes, Kambartel, Kaulbach, Veraart (Hamburg: Felix Meiner Verlag, 2nd edn., 1983), 139–140; also in *Posthumous Writings*, ed. Gabriel, Hermes, Kambartel, Kaulbach, Veraart; trans. Long and White, (Chicago: University of Chicago Press, 1979), 128–129. Ricketts claims that in the argument Frege relies on his elucidation of judgment as recognition of truth ('Logic and Truth in Frege', 129). He holds that Frege's notion of judgment applies only to truths and is thus 'quasi-factive'. He sees the argument as an attempt by Frege to show that the definability of truth would make judgment conceived in this factive way impossible. Ricketts does not note, however, that the argument never invokes the notion of judgment. The notions that Frege does use in the argument are manifestly non-factive ('decide whether a representation corresponds to actuality', 'compare representation and reality'). It is possible that Frege is assuming some connection between factive and non-factive judgment, but he nowhere makes this explicit. So I regard Ricketts's reading as implausible. It does not accord with the text. I think that Frege's argument is most charitably understood independently of any special notion of factive judgment.

[11] 'Logic and Truth in Frege', 131. I find Ricketts's account of the argument somewhat obscure. But I believe that I understand it well enough to make the following points. I also think that Frege's argument makes a charge of circularity, not of regress.

[12] I believe that Frege himself writes in a way that strongly suggests the mistake. Cf. 'The Thought', 354; in *Kleine Schriften*, ed. I. Angelelli (Hildesheim: Georg Olms, 1967), p. 345; O 61. I am not convinced, however, that Frege made the mistake.

Still, Frege *did believe* that truth is not a property, at least not a property of thoughts or sentences. More cautiously put: he was uncomfortable with this way of construing truth.[13] Although *property* is not a theoretical term that Frege uses in his theory, his theory is naturally read as not conceiving truth as a property.

Frege explicitly makes use of four expressions that are conceptually associated with truth. There is the horizontal—translated 'is the true'—which denotes a function from truth-values to truth-values.[14] There is, second, the singular expression 'the true', which denotes the extension of the concept or function denoted by the horizontal. Third, there is 'denotes' applied to sentences. This expression denotes a function from sentences, or their nominalizations, to truth-values. Fourth, there is the predicate 'determines' (or in Church's usage 'is a concept of') as applied to thoughts and components of thoughts. This expression denotes a function from thoughts to truth-values (more generally thoughts and thought components, or senses, to the denotations of expressions that express them).

Denotes the truth-value truth could perhaps be seen as a property of sentences. There would be a corresponding property of thoughts, *determines the truth-value truth*. But Frege clearly thinks that the fundamental notions are denotation (determination) and the truth-value, truth. The truth-value truth, which judgment 'moves toward', is an object, not a property. 'Denotes' is a two-place relational predicate under which sentences and truth-values, pairwise, fall. Analogously, for 'determines'.[15] Even allowing for Frege's loose, non-technical use of 'property', none of these expressions applies to or denotes properties of thoughts. It seems to me that this point about Frege is obvious. One does not need a strained interpretation, lacking textual backing, of the argument against the definability of truth to support it.

It is clear that Frege thought that the sense of the horizontal function sign does not add to the sense of declarative sentences to which it is attached. He seems to make a similar claim for the two-place predicate 'denotes' as applied

[13] 'The Thought', 354–355; *Kleine Schriften*, 345; O 61–62.

[14] Despite this translation, Frege clearly takes the horizontal to be more basic than the singular expression 'the true' that denotes the truth-value. Similarly, the concept denoted by the horizontal is more basic than the object which is the truth-value. Frege explains the latter in each group in terms of the former.

As mentioned before, the horizontal maps anything other than a truth-value onto falsehood.

[15] Church called this the *concept of* relation. As with Frege's notion of denotation, the concept-of (or determination) relation between a thought and its truth-value is a special case of the relation between any thought or thought component and what the corresponding linguistic expression denotes. It is to be understood, however, that this relation is a two-place relation. There is no argument place for a linguistic expression. Cf. Alonzo Church, 'The Logic of Sense and Denotation', in P. Henley, H. M. Kallen, and S. K. Langer (eds.), *Structure, Method and Meaning: Essays in Honor of Henry M. Sheffer* (New York: Liberal Arts Press, 1951). Frege is clearly aware of the possibility of using an expression for this relation. Cf. 'Logic' (1897), in *Posthumous Writings*, 229; in *Nachgelassene Schriften*, 140–141. If Frege had developed an intensional logic for attributions of belief, as Church did, I believe that he would have introduced the predicate 'determines' into his formal scientific language.

to a sentence and the true, and for the analogous two-place predicate that is applied to a thought and the true. He seems to regard all three predicates (or predicate applications) as having this sense-redundancy feature.[16]

Ricketts makes much of Frege's claims that truth is a goal of judgment, not a property of thoughts, and of Frege's further claim that 'it is true that ——' has the same sense as '——', for sentential fillings of the blank. Neither of these claims seems to me a high point in Frege's work. As I noted, the latter claim is surely false. There is no question that Frege made them, however. Ricketts draws strong conclusions from these claims that are intended to show that Frege could not have had a semantics—and consequently could not have had intuitive semantical notions like *logical validity* and *logical consequence*. I believe that here again the conclusions are not supported by textual evidence.

One conclusion is that although, according to the redundancy view, 'true' has a sense that does not add to the sense of sentences in 'it is true that ——', 'true' does not 'mean' (denote) anything.[17] I can find no evidence that Frege held this view. Ricketts does not present any. Frege clearly explains what concept the horizontal denotes. There are other passages in which Frege speaks of the denotation of 'true', without indicating that he does not mean what he says.[18]

Another of Ricketts's conclusions is that the relation between a thought and a truth-value is 'not describable in a sentence'.[19] This conclusion does not follow from either the redundancy view of the sense of 'it is true that' or from

[16] The three predicates are the horizontal, 'denotes', and 'determines' (as applied to a relation between thoughts and truth-values). Cf. 'Logic in Mathematics', in *Posthumous Writings*, 234; in *Nachgelassene Schriften*, 252. Frege is quite explicit in counting these latter two, meta-level expressions predicates. He indicates that they apply to relations not to properties.

[17] Ricketts, 'Logic and Truth in Frege', 134.

[18] e.g. 'The Thought', O 59.

[19] Ricketts, 'Logic and Truth in Frege', 135. Ricketts writes, 'The relationship between a thought and its truth-value is not describable in a sentence—it is not a matter of a thought's falling under a concept or of a relation's holding between two objects. Rather, the relationship between the thought that Socrates is mortal and the True is linguistically expressed by an indication of the asserting force with which a sentence expressing the thought is uttered by someone who has recognized-the-truth of the thought.' I believe that in his exposition of this conclusion, Ricketts runs together the relation between the thought and the truth-value with aiming at the truth-value in judgment, which carries the force of assertion or judgment. This is exactly what Frege inveighs against—except that Frege is more concerned with the mistake of assimilating judgmental force to the relation between a thought and its truth-value true than with the mistake of assimilating that relation to judgmental force. The relationship between a thought and its truth-value is very explicitly, in Frege's writing, a relation that holds between two objects (a thought and a truth-value). Frege nowhere states that the relationship between thought and truth-value is not describable in a sentence.

There is a sense in which judgmental force, on Frege's view, cannot be fully captured in any description or declarative sentence. Frege does make this claim. But he does not claim that the denotational relation cannot be described. As I have noted, he often describes the relation.

Rickett glosses Frege's word '*einzigartig*' (*sui generis*) as indicating that its content does not comfortably fit in any logical category (ibid. 132). No textual support for this gloss is given. I believe that what Frege says in the context of his use of the word does not support any such paraphrase.

Frege's views that truth is the goal for judgment and that truth is not a property of thoughts. Again, Ricketts does not provide the slightest textual support for attributing the conclusion to Frege. Frege quite explicitly describes the relation between a thought and its truth-value: '. . . there must be something associated with a sentence which is different from the thought, something to which it is essential that the parts of the sentence should have denotations. This is to be called the denotation of the sentence.'[20] The denotation's being associated with the thought (by being a denotation of the sentence that expresses the thought) is the described relation between thought and truth-value. Frege describes it in these two sentences. Clearly he could have described it in one! I know of no reason to think that Frege thought that the relation could not be described in a sentence in a formalized language—formalized within the *Begriffsschrift* with "non-logical" predicates. Moreover, it can be.[21]

It is well known that Frege claims that logic itself would be dispensable if our language were a perfect language. The language would exhibit, and operate in accord with, logical form and logical inference rules without any need to characterize the form, or state the inference rules separately. Since our language is not a perfect language, he indicates that we have a need for logic. He also implies that the word 'true' is indispensable in explaining logic, and even in 'laying the foundations' of logic.[22] It should be noticed that Frege

[20] "Introduction to Logic" (1906), in *Posthumous Writings* 194; *Nachgelassene Schriften*, 211.

[21] I know of no difficulty in principle in formulating a systematic type-theoretic semantics in the Fregean mode. A condition on such a semantics is that it be capable of proving formal analogs of

Sentence 'Fa' denotes the true if and only if the denotation of 'F' maps the denotation of 'a' onto the true

and

'Fa' denotes the true if and only if Fa.

One might have semantical rules for functions (extensionally conceived) like:

(F) ('is a horse' f-1-denotes F if and only if (x)(Fx if and only if x is a horse)).

'f-1-denotes' in this usage is a two-level two-place concept, taking objects (signs) and first-level functions into truth-values. 'o-denotes' as applied to singular (object-level) expressions would be a single-level two-place concept, taking a pair of objects (a sign and an object) into truth-values. Then the two 'denotes' (presumably together with further semantical predicates for the quantifiers) would have to be coordinated in order to prove the analogs of the Tarski biconditionals mentioned above.

A semantical theory for n-order function signs would be given in a theory that is at least $n + 1$ order. Frege himself thought that each function is equivalent to at most a function of level three. But this is incorrect, on any natural construal, by Cantor's theorem. (I am indebted to Terry Parsons on this point.) Although Frege may have hoped to collapse the type theory into a three-level theory, it appears that, as with almost any type theory, the semantics for a given level must always recede into a higher-level language. Still, there is no reason to think that on Fregean principles the semantical relations are not describable within the type-theoretic constraints.

[22] 'My Basic Logical Insights' (1915), *Posthumous Writings*, 252; *Nachgelassene Schriften*, 272. I believe that Frege had the notion expressed by the horizontal in mind in his discussion here. In 'On Sense and Denotation', 34 of the original, one can see the effects of his separating force from predication, in particular predication involved in the horizontal. He insists that force does not reside

explicitly states in this passage that 'true' is indispensable in logic. Its uses in logic would be dispensable only in a perfect language that had no need for a logic that explained logical structure and inference rules.

It is an open question just which term 'true' Frege is claiming would be dispensable, along with logic, if our language were perfect. It is likely that he has in mind the semantical predicates 'denotes' as applied to sentences and 'determines' as applied to thoughts. These notions are the ones used distinctively in the formulations of rules of inference. 'Is the true', the horizontal, is a primitive in Frege's *Begriffsschrift*. I think that the judgment stroke and the horizontal would be present, according to Frege, in a perfect-language formulation of any special science.

The conclusion that the relation between a thought and a truth-value is 'not describable in a sentence' might be encouraged by imagining a perfect language in which both logic and semantical occurrences of 'true' are dispensable and then reflecting on Frege's redundancy theory of the sense of 'true'. Frege appears to hold the redundancy-of-sense theory about all three 'truth predicates' in his system. One might reason as follows: ' "——" denotes the true' has the same sense as '——'. The predicate "denotes" would not be needed in the perfect language. So there is no sentence describing the relation between sentences and the truth-value, true, in a perfect language. Analogous reasoning could be extended to any analog of "denotes" that applies to the relation between thoughts and the truth-value, true.[23]

This reasoning would be unsound.

In the first place, Frege never says that 'true' is dispensable in a perfect language—only that logic is dispensable in a perfect language and truth is indispensable for logic. He does not claim that in a perfect language there would be no non-logical (special-science) uses for 'denotes' or its analog, 'determines'. So the second step is doubtful.

In the second place, since sentences '——' and ' "——" denotes the true' have, according to Frege, the same sense, the former sentence specifies the relation just as well as the latter. So the move from the second step to the conclusion is fallacious. I know of no reason to think that Frege would deny that in a perfect language either the denotation of 'the true' or the relation between thought and the true would be denoted, or made reference to. I will return to this point. There are several instances in Frege's work where he says that sentences with apparently disparate ontological commitments have the same sense. His view is that adequate understanding of the

in this predicate (as he has suggested in *Begriffsschrift*, section 3). His sense-redundancy view is directed at the idea that saying a thought is true, or saying 'it is true that——' captures the force of a judgment. In his formal work, he comes to regard 'is the true' as the fundamental object-level predicate that applies to gerundized construals of sentences. But he also makes use of the meta-level predicate in various ways, such as in his exposition of his axioms and in his statement of his inference rules. Cf. the last pages of section II of 'Frege on Truth' (Ch. 3 above).

[23] Such an extension is strongly suggested by Frege's remarks in 'On Sense and Denotation', O 34.

senses of the sentences enables one to think about each sentence's denotational commitments equally well through thinking a thought through the other.[24]

In the third place, Frege explicitly calls truth a predicate and indicates that it applies to a relation between a thought and its truth-value (when it is true). As I have indicated, he often describes the very relation that is supposed to be 'not describable'.[25] So the conclusion, in the third step of the argument, is surely not Frege's conclusion.

One further conclusion that Ricketts draws is that 'there can be, in a sense, no genuine theorizing about logic. There is only theorizing within logic—the proof of derived logical laws from basic logical laws and the application of logic in formal proofs within the framework of the begriffsschrift to the laws and facts uncovered by the special sciences.'[26] This conclusion seems to be associated with the claim, 'Frege's understanding of truth, in precluding a genuine truth-predicate—one usable in generalizations—also rules out truth-conditional semantics, rules out, that is, a theory of how (the thoughts expressed by) sentences are determined to be true or false by the items referred to in them'. Ricketts adds that the attempt to state truth conditions of the sentence 'Sea water is salty' 'yields, on Frege's view, only the tautology that sea water is salty if and only if sea water is salty'.[27]

It would have been well to apply *modus tollens* to this form of reasoning and find what is wrong with it rather than to accept the patently unacceptable

[24] One instance, associated with a corollary of Frege's Basic Law V is discussed in "Frege on Extensions of Concepts, From 1884 to 1903" (Ch. 7 below), section V of that article. Also see 'On Concept and Object', O 200; and 'On Sense and Denotation', O 34: 'These two objects [the truth-values] are recognized, if only implicitly, by everybody who judges something to be true . . . '. For further discussion of these issues, see 'Frege on Truth' (above).

[25] Cf. 'On Sense and Denotation', in the original, combine pp. 27 and 33–34, and especially p. 34. Also, 'Logic in Mathematics' (1914), in *Posthumous Writings*, 234; in *Nachgelassene Schriften*, 252 (where Frege counts the meta-linguistic 'denotes the true' a predicate); and 'Introduction to Logic' (1906), in *Posthumous Writings*, 194; in *Nachgelassene Schriften*, 211. Frege's point in denying that truth is a property is partly to insure that predications of truth of sentences or thoughts specify a relation between the sentences or thoughts and the true. Of course, judgment and its force are yet further matters.

[26] Ricketts, 'Logic and Truth in Frege', 136. Frege uses a semantical notion of truth in stating his inference rules. He believed that the truth predicate (cf. n. 25) that expresses this notion, like the horizontal (which is non-semantical), is indispensable in the practice of logic. There is clearly a kind of generalization involved in these uses of the truth predicate, 'denotes the true'. Since I have discussed the role of inference rules in his system in some detail in (Ch. 9 below) 'Frege on Knowing the Foundation', I will not enter into this issue further here.

[27] Ricketts, 'Logic and Truth in Frege', 140. In the last part of his paper, Ricketts advances a deflationary interpretation of Frege's ontological Platonism appealing to various ordinary-language points about judgment. This part of his paper makes little contact with Frege's texts. Moreover, it does not discuss the main passages where Frege gives his Platonism philosophical work to do—passages that I have discussed in 'Frege on Knowing the Third Realm' (Ch. 8 below). So I will not be discussing these claims. I should add that I think that the deflationary surrogates are, from a purely philosophical point of view, thoroughly inadequate to the task of accounting for a number of philosophical considerations that Frege was aware of.

conclusion that Frege's view of truth is in such clear contradiction with his development of a truth-conditional semantics.

Frege's semantics was not intended to justify the object-level or ground-level logic proper. It was intended to yield deeper understanding of it.[28] The *Begriffsschrift* was intended as a language that exhibits the fundamental justificatory relations between basic logical truths and arithmetical truths. Since the semantical elaborations are not justificatory, it is not surprising that Frege sees them as not part of a perfect language suitable for expressing basic logical truths and derivations from them that issue in arithmetical theorems.

Ricketts seems to be appealing to Frege's redundancy theory of the sense of 'true' to support his remarks about truth conditions being tautologies. 'Tautology' is Ricketts's term. I think that using it, at least in the standard modern way, to characterize Frege's thinking about such statements is anachronistic. The application also betrays a misunderstanding of Frege's conception of sense.

I have noted that Frege regarded applications of the semantical predicate 'denotes' to sentences and truth-values as having the same sense as the relevant sentences. He also applied his sense-redundancy theory to the analog of 'denotes' ('determines') that applies to the relation between thoughts and truth-values. Ricketts infers that Frege's sense-redundancy views commit him to holding that

'sea water is salty' denotes the true if and only if sea water is salty

has the same sense as

sea water is salty if and only if sea water is salty.

Or to take account of the fact that for Frege truth applies more properly to thoughts than to sentences, he would be committed to holding that the two sides of the biconditional in

the thought that sea water is salty determines the true if and only if seawater is salty

express the same sense.

I believe that up to this point Ricketts's reasoning is correct. It does not follow, however, that semantics is empty of scientific value in Frege's view. Before his acknowledgment of Russell's paradox, Frege regarded the two sides of the biconditional that constitutes his naive comprehension principle, a simple corollary of Basic Law V, *as expressing the same sense*.[29] He hardly regarded this corollary as a tautology, in any sense that would count it empty or without scientific value, or as not part of 'genuine' theorizing within logic.

[28] These points are discussed in my 'Frege on Knowing the Foundation'.
[29] Cf. n. 24. And section V of 'Frege on Extensions of Concepts, From 1884 to 1903' (Ch. 7 below).

The two sides of the truth-conditional biconditional, like the two sides of this naive comprehension principle, have the same cognitive value, according to Frege. How is it that they can be of scientific value?

At the beginning of 'On Sense and Denotation', Frege contrasts statements of the form 'a = a' with statements of the form 'a = b'. He says of the former that they hold apriori. He says that the latter 'often contain very valuable extensions of our knowledge'.[30] This may suggest that statements of the form 'a = a' do not contain valuable extensions of our knowledge. But Frege never says that they lack scientific value or usefulness. He simply says that statements of the two forms have different cognitive values and that statements of the 'a = a' form are apriori.

I believe that especially in cases where the sense of an expression has logical structure, Frege thought that there is room for informativeness or scientific value in expressions of the same thoughts in different ways. Sense is cognitive value. Cognitive value is what a fully rational, fully informed thinker, with complete mastery of the language, grasps.[31] Only a fully informed individual with complete logical understanding would find the semantical biconditionals or the corollary of Basic Law V (imagining that it had turned out to be a logical law) to be cognitively uninformative. Where there is incomplete understanding, expression of the same thoughts in different ways might well be cognitively valuable. Frege makes it clear that he thinks that logic is necessary only because we have an imperfect language. The same point would apply to the biconditional theorems in a semantics for logic.[32] This imperfect condition of ours is both cause and effect of incomplete understanding.

Frege develops this sort of point in his account of definitions. Definitions cannot add to knowledge in the most idealized sense of 'add'. The definiens has the same sense as the definiendum. If one has a perfect grip on the structure of a thought, then the definitions are dispensable in proof, and they are always, logically speaking, dispensable. But even definitions can be theoretically fruitful and of scientific value—especially if they express structure in a perspicuous way. For one's grip on the structure of a thought goes through language, and often a definition is needed by the reasoner to hold in mind the complex sense structure associated with a word, or to make clear what had been only imperfectly or incompletely grasped.[33]

[30] 'On Sense and Denotation', O 25.

[31] Cf. 'Frege on Extensions of Concepts' (Ch. 7 below), section VI; 'Frege on Sense and Linguistic Meaning' (Ch. 6 below), and the discussion of this point in the Introduction to these essays.

[32] There is the additional familiar point that although the Tarski-type biconditional theorems in a semantics seem trivial, the derivation of them from the semantics for the parts of sentences (or thoughts) is not at all trivial. I believe that Frege was aware of the analog of this point for his own systematic semantics.

[33] Frege, *The Foundations of Arithmetic*, section 70; 'Foundations of Geometry II', in *Collected Papers*, 300–301; *Kleine Schriften*, 289–290; O 302–303; 'Logic in Mathematics', in *Posthumous Writings*, 209, 216–222; *Nachgelassene Schriften*, 225–226, 234–240.

The fact that Frege regarded his naive comprehension principle as expressing the same senses on each side of its biconditional, and the fact that he regarded semantical truth-theoretic biconditionals in the same way, shows that he believed that more logical structure is involved in a sense or thought than is uncovered even by the linguistic structure of any one sentence in his logical notation. Thus the thought expressed by a quantified sentence which in Peano notation would be (x)(Fx ↔ Gx), is committed not only to first-level functions denoted by 'Fx' and 'Gx', but to courses of values associated with those functions. An ideography truly adequate to the thoughts being expressed would have to have an extra dimension, to account both for inferences associated with quantification into the function's argument place and for inferences associated with the identity of the courses of values.

Frege's remarks in his late 'My Basic Logical Insights' (1915) take on a deeper significance in the light of these points. Frege writes, 'If we had a logically perfect language, we would perhaps further need no logic or we could read it [logic] off the language. *But we are at a vast distance from being in this condition*'.[34] It appears that at least during the time when he was proposing Basic Law V, even his own ideography was not seen as one in which one could read off all relevant thought structure from any given sentence. Inferences from a thought involving extensions of concepts are not explicit in a sentence of simple quantification or predication. But both structures are determined in the sense or the thought. So more structure resides in a sense or thought than is uncovered by the linguistic structure of any given sentence even in the *Begriffsschrift*. A person who fully and deeply understands the right side of these biconditionals would fully and deeply understand the relational structure indicated in the left sides.

The same point applies to Frege's understanding of the science of semantics. He thought that the analogs of Tarski's biconditionals do express the same sense on each side of the biconditional. A person who fully understands the thoughts expressed by the biconditionals would learn nothing from them. In a perfect language, both structures would be mentioned. Frege regards logic as worthwhile given that languages are not perfect (do not in their form reveal all relevant logical inferences from the thoughts they express). He would, or certainly could, take the same view about semantics. It is scientifically worthwhile in explaining the foundations of logic, given that our languages are not perfect and given that incomplete understanding of thoughts expressed in our languages is common among users of the language—including users of logic.

It is important to recognize differences between Frege and his successors in logical theory. It is also important to see Frege's work in the broad perspective of the history of logic. Frege clearly initiated modern semantics. He had an unprecedented grasp of logical form and logical structure. He

[34] *Posthumous Writings*, 252; *Nachgelassene Schriften*, 272; the emphasis in the quotation is mine.

showed how the semantical structure associated with sentential parts (or thought components) are systematically associated with the truth-value of sentences or thoughts. His motive in producing this account lay in accounting for good deductive inference. He gave no central explanatory role to modality in his account. He certainly had and deepened the understanding of the intuitive notion of (formal) logical consequence. His conception of logic centered on the generality of its axioms and its subject-matter applications.

While it is certainly true that he did not give a model theory, did not raise the questions that led to the flowering of meta-logic (most notably in the completeness and incompleteness theorems), did not foresee the seriousness of the semantical paradoxes, and may have thought of semantics as more loosely related to logic than it in fact is, his conception of formal consequence, understood semantically, engendered the explosion of progress in semantics and meta-mathematics. His conception is in fundamental ways continuous with that progress.

Frege also differed from many of his successors in having superior insight into the ontological entanglements of logic and superior insight into its epistemology. He had a forthright conception of the dependence of the truths of logic on a subject matter. Logic for him is the most general science of being. This construal of his view is not less true in light of his type-theoretic understanding of being. In his view of logic as a science of being, he was joined by Russell and Gödel but by few others among his great successors. I believe that on this matter he had deeper insight than Wittgenstein, Carnap, Tarski, and even Quine. He also saw clearly that logical knowledge lies in understanding—fundamentally the kind of understanding that goes with competence in making good inferences and understanding logical truths. Little or no further warrant is to be gained through semantical work applied to logic. Such work clarifies and deepens understanding without fundamentally strengthening warrant.

Let me turn now to the fourth thing that I want to do in this postscript—speculate a little on Frege's views about semantics in a somewhat wider context. As I have noted, late in his career Frege states that 'true', presumably as formalized by the semantical predicates 'denotes' and 'determines', is indispensable in logic, but that logic is dispensable in a perfect language. I have noted that this view does not imply that the semantical predicates are dispensable in a perfect language. Frege uses his semantical predicates in expounding his inference rules in his 'extensional' logic (his logic that does not explicitly deal with senses). He would also have used the predicate 'determines' if he had elaborated the logic of oblique contexts—contexts of propositional-attitude attribution—that do specifically quantify over senses or thought contents.

I raise three questions. First, does Frege's claim that logic is dispensable in a perfect language apply only to the meta-logical elements of the logic—the explication of basic logical terms, the elaboration of the language's syntax and inference rules—or does it apply also to the object-language axioms of the logic? Second, if Frege had elaborated a logic for oblique contexts, would he have maintained that the axioms governing senses and their relations to their *determinata* would occur in a perfect language—or would he have regarded such axioms as dispensable? Third, are there non-logical uses for the semantical predicates that Frege would have admitted in a perfect language?

As to the first question, I think that there is reason to believe that when he maintained that logic is dispensable in a perfect language, Frege had in mind only the meta-logical elements of logic—the explication of terms, the elaboration of syntax, the semantics of logical symbols, and the formulation of inference rules. His idea was that all meta-logical explanation would be unnecessary in a perfect language because these matters could be "read off" the language's structure and meaning. One would understand the terms, structure, and valid inferences simply by allowing the perfect language to inform one's thinking. The object-language axioms of logic would, however, remain. They state the most general truths in a science of being. Even though they could be "read off" a perfect language by someone with perfect understanding, there is no reason to dispense with them in a perfect language that serves science. They do not "teach" the language as the meta-logical parts of logic do. They are fundamental truths about all subject matters.

These points strongly support an answer to the second question. In Frege's "extensional" logic—the one that he uses to formalize mathematics—there is no need to make any reference to senses. The only entities that one makes reference to are functions and extensions (including truth-values). In a logic that deals with oblique contexts, one that would be needed in accounting for valid inference in attributions of propositional attitudes, the axioms of the logic must make reference to senses or modes of presentation. They must also make reference to the determination relation between modes of presentation and their *determinata*. Thus predicates for mode-of-presentation contents and for the determination relation would not be "non-logical". In the object-language axioms, there need be no reference to linguistic symbols. But there must be reference to determination relations between senses (modes of presentation) and functions or extensions. Frege would have thought that in a perfect language, there would be no need for meta-logical formulations of inference rules and no need for semantical statements of the relation of expression between *symbols* and senses. But I believe that he would have regarded the object-language axioms of a logic for propositional attitudes, which must make reference to senses (or thought components) and their determination relations to denotations, as indispensable in a perfect language.

Axioms of logic are principles in a science—the most general science. A perfect language serves science. I think that Frege thought that some thoughts (thought contents) are, like the numbers, implicitly present in every subject matter. Hence they would be a proper subject for a universal science of being. Frege's Platonism implies that thoughts that are logical truths (including arithmetical thoughts) are in a sense implicit in any discourse, no matter what the topic. What this means is that no matter what subject matter a piece of language is used to discuss, that piece of language contains expressions that are related (by the expression relation) to logical truths. In this sense, thoughts that are logical truths are just as ubiquitously related to any discourse as numbers or truth-values. A universal science of being should make reference to them. Of course, unlike numbers and truth-values, they are not part of the (denoted or determined) subject matter in the discourse of a language formulated in extensional logic, one that formalizes no oblique contexts from natural language. But a logic that governs the relations between thought contents, on one hand, and truth-values and other entities such as functions, on the other, will contain axioms that are logical truths. The logic will include an impersonal conception of assertion or judgment, idealized to apply to the theorems of logic—*it is to be rationally and ideally judged that*_______ (where the blank is to be filled by axioms or theorems of logic).[35] Such a logic of thought contents would be applicable whether or not anyone ever actually had propositional attitudes. It would be applicable even in discourses that do not explicitly discuss propositional attitudes (discourses that could at most be part of a special science). In an "intensional" logic that formalized such sentences, thoughts would be part of the denoted subject matter of the language. They would not be implicit in a subject matter merely by being expressed. Such a subject matter would, from Frege's Platonic point of view, be just as relevant to and present in any subject matter as the truth-values and the numbers. The thought contents and truths quantified over in the object-language axioms of such a logic would be subject matter in a universal science of being. The relevant science would include an account of the relation between thought contents and their subject matters (pre-eminently thoughts and truth-values) *in its object language*. Such axioms would not be dispensable in a perfect language. Of course, we now know that things cannot be so simple. But Frege was not very sensitive to the threat of semantical paradox, and related paradoxes.

I believe that the answer to the third question is straightforward. There is every reason to think that Frege would have admitted special-science uses for the semantical predicates—predicates governing the relations between *symbols* and senses/denotations. There is every reason to think that he would have acknowledged that semantical predicates can be incorporated into a special science. There is no evidence at all that Frege took relations between language, including empirical language, and its senses and denotations to be

[35] In such a language, Frege would still not regard the force of judgment as being captured by the idealized propositional-attitude verb. The vertical judgment stroke would still be needed. No predicate would formalize it.

anything other than factual. A special science would have to deal with such relations. Such a science could deal with thought contents that are clearly not part of every subject matter, even implicitly. Such a science can be formalized within the *Begriffsschrift*.

4 *Frege and the Hierarchy (1979)*

At the level of surface syntax in statements of propositional attitude, certain classical principles of substitutivity commonly associated with extensionality fail. The principles are

$t = s \rightarrow (A^x/t \leftrightarrow A^x/s)$

and

$(A \leftrightarrow B) \rightarrow (\ldots A \ldots \leftrightarrow \ldots B \ldots)$

where t and s are terms; A and B are formulae; 'A^x/t' signifies the result of substituting t for variable x in one or more occurrences; and '$\ldots A \ldots$' signifies any formula in which A occurs. These principles fail at the surface level in that, for example, from 'Bonn is the birthplace of Beethoven' and 'Anton believes that Bonn is in Germany' (on one reading), we cannot deduce 'Anton believes that the birthplace of Beethoven is in Germany'. Analogously, from 'Berg was Austrian and Brahms was German', the sentence, 'Anyone who believes that Brahms was German believes that Berg was Austrian' does not follow as a matter of logic.

The most thoroughly studied response to such failures of substitution (both in belief contexts and elsewhere) is to give up the principle of extensionality: the principle that the denotation, or extension, of an expression (including the truth-value of a sentence) is a function of the denotations, or extensions, of its semantically relevant parts as they occur in the expression. The response involves taking terms like 'believes' (or 'believes that') to be sentential operators, restricting the classical syntactical laws of substitution and quantification, and relativizing the semantical relation of denotation (or satisfaction) to a possible world. Let us call this 'the sentential operator approach'. This general approach to substitution failures in natural language has unquestionably deepened our understanding of semantical structure. But for many purposes, the key notion of possible world seems less clear than the discourse it is introduced to interpret. Moreover, it is hard to take seriously the notion at face value, as not to be explained in more primitive

I am indebted to Nathan Salmon for comments on an earlier version, and to the National Endowment for the Humanities for its support.

terms, since its typical metalinguistic use demands that there be worlds other than the actual one and individuals other than the actual ones. The appeal to possible worlds is more naturally taken as a heuristic aid in constructing completeness and consistency proofs and in formal reasoning with modal statements—a half-way house in arriving at an intended interpretation of the discourse being studied. Moreover, as applied to statements of propositional attitude, the notion of possible world has less intuitive appeal than elsewhere.

A further problem with the sentential operator approach is that it tends to leave one without the resources to capture certain natural inferences. For example, from

(1) Arnold believed that dominant resolves to tonic

and

(2) The most basic point of classical harmony is that dominant resolves to tonic,

we may conclude

(3) Arnold believed the most basic point of classical harmony.

(3) illustrates the problem most clearly. Since the sentential operator approach takes expressions like 'believes' (or 'believes that') to apply syntactically to sentences, it cannot easily account for the noun phrase that typically follows such expressions and for inferences (either substitutions or generalizations) that turn on that phrase.

The problem about generalization runs deeper. An operator approach could hold that sentences of the form 'Arnold believed something' are to be represented with a substitutional existential quantifier ['($\exists p$) Arnold believed that *p*'], which takes sentences as substituends. But such a representation assumes that all of Arnold's beliefs are expressible in English sentences. This assumption seems to me implausible for the general case, but I shall not argue the point here. Suffice it to say that unless the assumption is sound, the operator approach does not have even the beginning of an account of the quantification onto the content clause.

Frege took another approach to the substitution failures. Instead of treating them as counter-examples to extensional principles, he regarded them as evidence that the occurrences of terms and sentences within that-clauses (and other such contexts) have different semantical functions than occurrences of those same terms and sentences outside of that-clauses. Thus 'Bonn' in the report of Anton's belief denotes not the German capital, but something else—what Frege called its oblique denotation. And the expression 'Brahms was German', as it occurs in the relevant that-clause, functions not as a sentence that is true or false, but as a complex singular term denoting (obliquely) an object of possible belief. Since 'believes' is, on the Fregean view, a predicate that applies syntactically to terms, there is no difficulty in accounting for inferences like (1)–(3). What is more, the view avoids the unintuitive elements of the possible world analysis.

My purpose in this paper is to discuss from a rather abstract point of view the effect of embedded oblique contexts on the Fregean strategy. I shall focus on the question: Do embedded oblique contexts pressure one to postulate an infinite (or indefinitely high) hierarchy of entities, with each level of the hierarchy serving to provide denotations for expressions in different degrees of embeddedness?

I

Two ways of formally accounting for the shifts of denotation by and within that-clauses are implicit in Frege's work. According to one (Method I), the formal representation allows expressions to have different denotations according to the syntactic context in which they occur. In one sense, this view allows ambiguity into the formal representation itself: a given expression has various denotations. But the view does not violate any fundamental assumption about the formality of formal systems as long as the denotation of an atomic expression can always be rigorously specified by reference to the syntactical form of complex expressions in which it occurs.

On Method I, then, the relation between an expression and its denotation is relativized to a syntactic context.[1] Thus, roughly speaking, a semantical theory for natural language would contain rules like: the denotation of 'Bonn' relative to its occurrence in any context of class T is, if anything, Bonn—where T is a rigorously defined class of syntactic contexts (what are sometimes called 'transparent contexts'). The denotation of 'Bonn' relative to ordinary oblique that-clause contexts would be something other than Bonn. Method I requires that the classical laws of substitutivity be restricted. On the assumption that Bonn is the capital of Germany, we may interchange 'Bonn' and 'the capital of Germany' only in syntactic contexts where the denotation of the singular terms is the same as it is in identity contexts. Despite this restriction on the syntactical rule most closely associated with extensionality, the account that results from following this method is fully extensional: The truth-value of a belief sentence is determined solely by the denotations (or satisfaction values, or extensions) of its parts—as they occur in the sentence.

The syntactical notion of substitutivity and the semantical notion of extensionality are sometimes lumped together. And often the former is used as

[1] Method I appears to be suggested in Frege, 'On Sense and Reference', in P. Geach and M. Black (eds.), *Translations from the Philosophical Writings of Gottlob Frege* (Blackwell's, Oxford, 1966). It is perhaps the most straightforward implementation of Frege's slogan: "Never ask for the meaning (*Bedeutung*) of a word in isolation, but only in the context of a sentence" (*Foundations of Arithmetic*, trans. J. L. Austin (Evanston, Ill. Northwestern University Press, 1968) p. x).

criterion for the latter. But Frege's strategy (assuming for the moment that he followed Method I) is illuminatingly described by distinguishing them. Frege's strategy was to treat apparent counter-examples to extensionality as cases of ambiguity. His primary tool for the analysis of language was to treat the truth-value of a sentence as a function of the entities denoted by the semantically relevant components of the expression. The substitution failures showed only that expressions in non-oblique contexts (say, identity contexts) denoted something different from what they denote in oblique contexts. The motivating principle of the analysis is preserved on Method I, even though the denotation of an expression can be properly specified only in the context of a containing sentence.

There is no reason to think that Frege would have held that languages involving the relevant ambiguity cannot be extensional. There is no reason for thinking that Frege would have held that extensionality fails in natural languages and can be preserved only in 'well-constructed' formalized languages where the systematic ambiguity that he attributed to natural languages is removed. Such a thought would put the matter in the wrong light: It is substitutivity that, according to Frege, fails in natural languages. The intuitive notion of extensionality, which he took to be fundamental to semantical evaluation, is unaffected. To allow for ambiguity, we must formulate the intuitive notion so that the relevant denotations or extensions of sentential components are understood to be the denotations or extensions of the components *as they occur in the relevant sentence*. But that qualification is hardly controversial. Modal logic and current belief logics, of course, flout both substitutivity and extensionality principles under their usual interpretations. So Method I contrasts with these approaches semantically, if not in rejection of the principles of substitutivity.

The second Fregean method for explicating the shifts of denotation (Method II) is to represent expressions in natural-language oblique contexts with symbols which are different from the symbols that represent those same expressions as they occur in ordinary contexts.[2] There is no need on Method II to relativize denotation to a syntactic context and no need to restrict the

[2] Michael Dummett, *Frege: Philosophy of Language* (London: Duckworth, 1973), ch. 9, suggests that Method I is clearly the method favored by Frege. But this is misleading. Frege's general insistence on avoiding ambiguity in a well-constructed language (perhaps the representing language) and a passage in a letter to Russell, December 28, 1902, militate against Dummett's suggestion: "Eigentlich musste man ja, um Zweideutigkeit zu vermeiden, in ungerader Rede besondere Zeichen haben, deren Zusammenhang aber mit den entsprechenden in gerader Rede leicht erkennbar wäre" ("Actually, in order to avoid ambiguity, one must in indirect speech have special signs whose relation to the corresponding signs in direct speech would be easily recognizable."), quoted in James M. Bartlett, *Funktion und Gegenstand* (Munich: M. Weiss, 1961), 19; also in John Wallace, *Philosophical Grammar* (Dissertation, Stanford, 1970), pp. 105–6. The chief proponent of Method II is Alonzo Church, 'A Formulation of the Logic of Sense and Denotation' in P. Henle, H. M. Kallen, and S. K. Langer (eds.), *Structure, Method, and Meaning* (New York: Liberal Arts Press, 1951); 'A Revised Formulation of the Logic of Sense and Denotation', *Nous*, 7 (1973), 24–33; 8 (1974), 135–156.

classical laws of substitutivity and quantification.[3] Method II again preserves the principle of extensionality. On either method the substitution failures in natural language are no more counter-examples to extensionality than fallacies of equivocation are counter-examples to modus ponens. We should not think of Frege as trying to translate non-extensional natural language into an extensional formal language (which is how Carnap and his students viewed the situation). Rather we should see him as arguing that, though ambiguous and syntactically misleading, natural language is covertly extensional.

In discussion of the two methods, it is often suggested that the first accounts for natural language whereas the second reforms it. This view is not clearly justified. One may regard the two methods as competing accounts of ambiguity in natural-language surface sentences. A standard device used in linguistic accounts of ambiguity is to subscript in the formal representation different representations of an ambiguous surface expression. Method II simply employs that device. What is partly at issue between the two methods is whether the ambiguity involved in the substitution failures should be resolved purely by reference to syntactic context, or whether it should be explicated, at least partly, by distinguishing the readings of particular expressions.

Interestingly enough, if Method I were applied directly to the surface sentences of natural language, it would be doomed. One cannot always determine purely on the basis of surface syntax whether or not an expression has an oblique denotation. In the sentence

(4) Schumann believed the Polish youth to be a genius

the expression 'the Polish youth' may be taken to have either an ordinary or an oblique denotation. Thus we may understand the sentence either as admitting or as blocking substitutions of other singular descriptions of Chopin for 'the Polish youth'. Method I is therefore best regarded as a strategy for constructing and explicating *formal representations* of surface sentences. As such, it will be committed to distinguishing the different interpretations of (4) with syntactically distinct (as opposed to merely *lexically distinct*) formalizaions. Now Method II is not thus committed. But I believe that the different interpretations of (4) *are* structurally distinct—and should be so construed by either Method.[4] Moreover, either Method should be able rigorously to define among the formal representations of natural-language sentences the syntactic contexts that represent oblique occurrences in surface

[3] If we wished to apply semantical analysis to the natural-language surface sentences, then on Method II, we would have to relativize the chief semantical relation to a formal reading (e.g.: 'Bank' relative to the reading '(Money)bank' denotes (Money)banks). The laws of substitutivity may be expected to remain unrestricted, since they work essentially on formal representations in any case. In general, it is easier to apply the truth predicate to canonical or formal readings of the surface sentences.

[4] For examples of representations that give syntactically distinct treatments to the different readings of the sentence, see John Wallace, 'Belief and Satisfaction', *Nous*, 6 (1972), 85–95; David Kaplan 'Quantifying In', in D. Davidson and J. Hintikka (eds.), *Words and Objections* (Dordrecht: D. Reidel, 1969); and my 'Belief *De Re*', *The Journal of Philosophy*, 74 (1977), 338–362.

sentences. It thus appears that Method II accounts for the shifts of denotation both by syntactical distinctions and by lexical distinctions. Is this not a sort of explanatory overkill?

The answer depends partly on whether Method II can offer compensatory advantages. It can. We have already noted that it provides a simpler semantical analysis for formal representations (it need not relativize the denotation or satisfaction relation to a syntactic context) and a simpler set of principles to account for deductive inferences (substitutivity and quantification laws need not be restricted).

On Method I even the grammatical categories should be relativized to a syntactic context. Thus Method II's investment in lexical complexity in accounting for ambiguity pays dividends in structural simplicity. I shall not try to judge the relative merits of this trade-off between the two methods. But I will later argue that from one point of view, Method I's lexical investment is no less substantial than that of Method II.

So far I have avoided asking after the nature of the oblique denotations of terms and sentences in natural language. This issue has been something of a sore spot for the Fregean approach. Frege himself believed that the oblique denotations are senses that terms and sentences express when they are used in ordinary, non-oblique contexts. And he regarded these senses as abstract entities, existing and applying to their associated denotations independently of any language that expresses them.

A major source of unwillingness to accept the overall Fregean approach has been refusal to think of senses as so completely independent of intentional agents. I approve of this refusal. But the Fregean semantical viewpoint does not of itself commit one to the Fregean ontology. One might follow Frege's general strategy for saving extensionality by appeal to shifts of denotation in natural language surface sentences (appealing to either Method I or Method II), but hold that the oblique denotation of expressions are the expressions themselves, not extra-linguistic senses.[5] And there are other alternatives. I mention the ontological question not because I want to pursue it here, but because I want to emphasize that Frege's semantical strategy has a general interest regardless of one's ontological viewpoint.

II

As we have described it, Method II holds that in unembedded oblique contexts like 'Bela believes Opus 132 is a masterpiece', the sentential

[5] I discuss one objection (the Church–Langford translation test) to taking expressions as oblique referents in 'Self-Reference and Translation', in M. Guenthner-Reutter and F. Guenthner (eds.), *Meaning and Translation* (London: Duckworth, 1978). For a systematic exploration of several aspects of the formal relations between the syntactical and Fregean ontologies, see David Kaplan, *Foundations of Intensional Logic* (Dissertation, UCLA, 1964).

expression 'Opus 132 is a masterpiece' is to be formally represented not by a sentence but by a term, say 'α'. For the sake of argument, let us suppose that 'α' denotes the proposition that Opus 132 is a masterpiece. According to Method II, 'α' may be exchanged for any other term (in the language of the formal representation) that denotes the same proposition. This yields reasonable results in representations of sentences containing unembedded occurrences of 'believes'. For suppose that 'π' represents 'Sergei's favorite proposition' and that Sergei is supremely enamored of the proposition that Opus 132 is a masterpiece. Then we can conclude: Bela believes α if and only if he believes π.

But embedded contexts present a new situation. Indeed, it can be argued on certain plausible assumptions that Method II must appeal to a hierarchy of some kind in order to represent embedded oblique contexts. The argument reduces to absurdity the view that 'α' may represent 'Opus 132 is a masterpiece' as it occurs in (5):

(5) Igor believes Bela believes Opus 132 is a masterpiece.[6]

Let us assume then that 'α' (which denotes the proposition that Opus 132 is a masterpiece) represents 'Opus 132 is a masterpiece' as it occurs in (5). (5), we shall assume, asserts a relation of belief between Igor and the proposition that Bela believes that Opus 132 is a masterpiece. I shall denote this proposition by the expression '$\Gamma_1(\beta_1, \alpha)$'. Thus on our assumptions, (5) is formalized as

(6) Believes (Igor, $\Gamma_1(\beta_1,\alpha)$)

By the principle of extensionality, the denotation of '$\Gamma_1(\beta_1, \alpha)$' is a function of the denotations or extensions of its parts. I shall assume that 'β_1' denotes the sense of 'Bela' and that 'Γ_1' denotes the sense of 'believes'—a function from β_1 and α to the relevant proposition.[7]

We assume the principle that a given sense is associated with a unique denotation or extension. Thus the proposition $\Gamma_1(\beta_1, \alpha)$ is associated with (or, in Church's terminology, is a concept of) a unique denotation or extension, its truth-value.

We assume that this truth-value is a function of the unique denotations or extensions associated respectively with the senses that determine the proposition. To put this another way, the truth-value of the proposition is a function of the denotations or extensions of expressions that express its component senses. Let 'β' express β_1 and denote Bela; let 'Γ' express Γ_1 and denote what

[6] In what follows I shall confine myself to embedded contexts that do not contain semantical terms like 'true', or terms whose use carries semantical implications, like 'necessary' or 'knows'. These expressions complicate representations of embedded contexts in ways that I prefer to consider separately.

[7] Already we have an inconvenience. In representing the second occurrence of 'believes' in 'Igor believes Bela believes Zoltan's favorite proposition', 'Believes' would most naturally be used to denote a two-place function from the sense of a term and the sense of a term to a proposition. But in this context, it has senses of sentences in its domain. But worse is to come.

'believes' denotes (or have its extension). Let 'α_0' express α and denote its truth-value. (We suppose that truth-value to be truth). Then '$\Gamma(\beta, \alpha_0)$' express $\Gamma_1(\beta_1, \alpha)$ and denotes its truth-value. 'Believes' originally applied to persons and propositions. But on our assumptions it has come also to apply to persons and truth-values. This leads to absurdity in short order.

For given the classical substitution laws of Method II, we may substitute any expression that denotes truth for 'α_0' in '$\Gamma(\beta, \alpha_0)$' and preserve the truth-value of '$\Gamma(\beta, \alpha_0)$'. (We speak of both sentences and propositions as having truth-value.) But '$\Gamma(\beta, \alpha_0)$' supposedly expresses the proposition that Bela believes Opus 132 is a masterpiece. So it seems to follow that if Bela believes Opus 132 is a masterpiece, he believes every truth.

The argument shows that on these assumptions 'Opus 132 is a masterpiece' in (5) cannot be represented by a term 'α' denoting the proposition that Opus 132 is a masterpiece. It is prima facie plausible to assume with Frege that the expression as it occurs in (5) should be represented by a term denoting the sense of the expression that represents 'Opus 132 is a masterpiece' as it occurs in unembedded belief contexts. Given this assumption, the argument can be replicated to show that the sentential expression as it occurs in doubly embedded oblique contexts must be represented by yet another term—and so on.[8]

Let us review the assumptions of the argument. The assumptions of extensionality and classical substitutivity are basic to Method II. Distaste for the hierarchy led Carnap to relinquish both.[9] I shall not discuss this large issue here except to register the view that no non-extensional theory of belief has provided an alternative that is plausibly superior to the Fregean approach. We shall discuss giving up classical substitutivity, but not extensionality, when we come to Method I.

I assumed that β_1 is the sense of 'Bela', that Γ_1 is the sense of 'believes', and that these senses are denoted in (5). Taking Γ_1 to be a function from senses to propositions is not strictly necessary to the argument. It is possible that one would want the sense of 'believes' to be individuated more finely than identifying it with such a function would allow. Moreover it is possible that one would want to represent the sense of 'Bela' not as an individual concept, but (in some more Russellian way) as some sort of function with propositions in its range. Such variations will not affect the argument as long as the other assumptions stand.

[8] The preceding argument is nowhere, to my knowledge, explicitly stated. Its conclusion and premises (near enough) are assumed in Church, 'A Formulation of the Logic of Sense and Denotation', and Kaplan, *Foundations for Intensional Logic*. I should note that the appeal to a hierarchy of senses in the Fregean system is motivated not only by embedded oblique contexts, but also by higher-level extensions of the 'paradox' of identity.

[9] Rudolf Carnap, *Meaning and Necessity*, 2nd edn. (Chicago: University of Chicago Press, 1956), 129–144, 232.

We have been accepting for the sake of exposition the Fregean assumption that unembedded oblique sentential expressions denote a proposition, and that oblique expressions generally denote senses. But the argument for the hierarchy does not depend on commitment to Fregean senses. For the relation between a linguistic expression and the sense it expresses, we could substitute a relation between the same expression and a standard syntactical name of it. For the relation between the sense of an expression and the denotation associated with it, we may substitute the relation between the standard name of an expression and the denotation of its denotation. The resulting argument would be equally plausible. And other ontological alternatives are open.

I assumed that the truth-value associated with a proposition is a function of the denotations or extensions of the expressions that express its component senses. This assumption is almost the principle of extensionality again, but it adds that senses are associated with the denotations of expressions that express them. The addition is, I think, hardly controversial. It may be thought to be if certain distinctions are not kept in mind. It is tempting to note that (5) may be read as follows:

(7) Igor believes of the proposition that Opus 132 is a masterpiece that Bela believes it.

(6) might be regarded as a first approximation analysis of this reading. Now one might go on to reason that the truth-value of what Igor believes is a function of the ordinary denotations of 'believes' ('Γ'), 'Bela' ('β'), and 'the proposition that Opus 132 is a masterpiece' ('α')—*not 'Opus 132 is a masterpiece* ('α_0'). In general, to find the truth-value of a proposition mentioned in a belief sentence that involves transparent contexts, we do not 'drop down' a type level for each of the component senses. Sometimes, as Russell emphasized, the components of such propositions will not even *be* senses: they may be individuals. None of these considerations, however, affects the argument. For (7) is not the relevant interpretation of (5). We may substitute coextensive phrases for 'the proposition that Opus 132 is a masterpiece' in (7) and preserve truth-value. But on one reading of (5), analogous substitutions fail.[10] It is to this reading that the assumption applies.

Finally, I assumed that for a given sense there is associated a unique denotation. This principle has been rejected in application to certain forms of language—for example, demonstratives. But these forms of language are

[10] Wallace, who presupposes Method II, attempts in 'Belief and Satisfaction', to avoid a hierarchy. The idea briefly is to treat (5) as one would treat (7). For Wallace a singular term denoting a first-level proposition or propositional function always represents a sentential expression in oblique contexts regardless of whether the context is embedded. The attempt to avoid the hierarchy fails for the reason stated above: (7) is not the relevant reading of (5). Another approach originally developed by Wallace in *Philosophical Grammar*, is held by Wallace to avoid the hierarchy and is so advertised by Davidson in 'On Saying That', *Synthese*, 19 (1969), 136–137. This approach is not tested in these passages on embedded contexts. If it is, it will be seen to be subject to our argument.

not ubiquitous. It is easy to present sentences which do not contain them. A thorough exploration of a semantics that rejects the principle generally has not, to my knowledge, been undertaken. In summary, the argument for a hierarchy on Method II seems very powerful.

It would be mistaken to try to simplify the argument so as to dispense with all assumptions except that of substitutivity, as follows. If (5) is formalized as (6), then if α is Zoltan's favorite proposition (which latter phrase we represent by 'π'), then (8) follows from (6):

(8) Believes (Igor, $\Gamma_1(\beta_1, \pi)$).

But on the assumption that (6) formalizes (5), (8) would seem to formalize

(9) Igor believes that Bela believes Zoltan's favorite proposition.

But (5) and the relevant identity do not yield (9)—Igor may never have heard of Zoltan. So (concludes the argument) on the assumption of classical substitutivity, 'α' cannot represent 'Opus 132 is a masterpiece' as it occurs in (5).

The argument does not work. (8) does not formalize (9). For since 'Zoltan's favorite proposition' is in an oblique belief context (at least the oblique context of the first 'believes'), it must denote its ordinary sense. (In any case the argument would prove too much, for it could be adapted to show that *no* reasonable representation could be maintained if classical substitutivity were preserved.) So to get the argument for the hierarchy, it seems that the longer route is necessary.

An infinite hierarchy of entities is, I suppose, theoretically unappealing, other things being equal. But the present hierarchy has considerably more intuitive content than has sometimes been imagined. We can bring this out as follows. According to (5), Igor has a belief that makes reference to a certain belief content (to the belief that Opus 132 is a masterpiece). As we ascribe it in (5), Igor's belief makes reference to that belief content in a certain way—it specifies it by applying a standard name. Igor might have had a belief that is best represented as making reference to that belief content in other ways, say by describing it as the first point made in a given chapter of a given book on 19th century chamber music. As we specify Igor's belief in (5), we make reference not directly (or not merely) to the belief content he ascribes to Bela, but to a certain means of ascribing it, a means that we ascribe to Igor. One means may be available to Igor—another, not. The subject matter of the hierarchy is thus certain means of ascription, means of ascribing those means, and so on.

III

Let us now view (5) through the lens of Method I. As conceived by Frege, that method too involves an infinite hierarchy of senses. There is not, however, the same pressure on Method I as on Method II to generate a hierarchy: The argument for the hierarchy that we gave in section II depended on the ability to substitute extensionally equivalent expressions in all contexts. Indeed, it has been suggested that the hierarchy is superfluous given Fregean principles and Method I.[11] Nevertheless, there is a sense in which the hierarchy is inevitable (given certain further assumptions) even on Method I.

We have described Method I as a strategy for creating a language L_I that formally represents the logical form of English. Our remarks about the semantical functions of expressions of L_I occur in a metalanguage ML_I. In order to give a systematic account of the truth conditions of sentences of L_I, we should formalize ML_I and give within it a theory of truth for L_I. If we are to give a theory that works on the structure of the sentences of L_I, we cannot construct ML_I on the plan of Method I. For to explain the denotations of complex expressions in terms of the denotations of their parts, we must be able to exchange within (translations of) the expressions of L_I different terms with the same denotations. On Method I, substitutions within expressions representing ordinary oblique contexts are allowable only if the expressions have the same sense. But the sort of expressions that need to be exchanged—e.g. γ, the denotation of γ relative to contexts 0, and the sense of γ relative to contexts *T*—clearly do not have the same sense. So on Method I the substitutions could not be carried out in the metalanguage, nor therefore could a systematic theory of truth be given for L_I. Thus ML_I has to be constructed so as to allow substitutions of extensionally equivalent expressions which do not express the same sense.[12]

Now suppose, contrary to Frege, that L_I lacks a hierarchy. Expressions representing transparent contexts denote their ordinary denotation and express their ordinary sense; expressions representing oblique contexts denote their ordinary sense and express no further sense. (Cf. note 11.) We shall state the semantical rules of ML_I relevant to accounting for the truth conditions of (5), as represented in L_I:

(a) $(x)(x = \mathrm{A}(\alpha, \ulcorner \mathrm{Igor_L} \urcorner, \mathrm{T}) \leftrightarrow x = \mathrm{Igor})$

(b) $\mathrm{Sat}(\alpha, \ulcorner \mathrm{Believes_L}\ (e_1, e_2) \urcorner, \mathrm{T}) \leftrightarrow \mathrm{Believes}\ (\mathrm{A}(\alpha, e_1, \mathrm{T}), \mathrm{A}(\alpha, e_2, \mathrm{O}))$

(c) $\mathrm{A}(\alpha, \ulcorner \mathrm{Bela_L} \urcorner, \mathrm{O}) = \mathrm{Sense}\ (\ulcorner \mathrm{Bela_L} \urcorner)$

[11] This language is proposed by Dummett, *Frege*, 266–269, as a means of avoiding a hierarchy. Dummett does not, however, investigate formalization of the metalanguage, which I shall argue urges a hierarchy at least from a certain point of view. Dummett's proposed language is not as compatible with fundamental Fregean principles as Dummett implies. In embedded oblique contexts, it controverts the principle that the denotation associated with a proposition is a function of the denotations associated with the component senses of the proposition.

[12] Actually, the argument I am giving does not require that ML_I have the *classical* substitution laws, though I shall assume here that it does. It might—as far as my argument goes—allow in certain contexts substitution only of 'necessarily' or 'logically' equivalent expressions *which differ in sense*, as long as the language is extensional in its semantics in the sense that the argument involving sentence (3) required.

(d) $A(\alpha, \ulcorner \text{Believes } (e_1, e_2)\urcorner, O) = C_3$ (Sense ($\ulcorner \text{Believes}_L \urcorner$), $A(\alpha, e_1, O)$, $A(\alpha, e_2, O)$)

(e) $A(\alpha, \ulcorner \text{Op. } 132_L \urcorner, O) =$ Sense ($\ulcorner \text{Op. } 132_L \urcorner$)

(f) $A(\alpha, \ulcorner \text{Masterpiece}_L(e_1)\urcorner, O) = C_2$ (Sense ($\ulcorner \text{Masterpiece}_L \urcorner$), $A(\alpha, e_1, O)$)

'$A(\alpha, e_1, O)$' signifies the assignment of any sequence α to expression e_1 as it occurs in any member of the class of oblique-representing contexts O. 'T' specifies the class of transparent-representing contexts. 'Sat' is the satisfaction predicate for L_I. I have subscripted mentioned expressions of L_I with an 'L' to distinguish them from the used expressions of ML_I. Thus 'Believes_L' of L_I (whose second argument expression may be exchanged only for expressions expressing the same sense) is to be distinguished from 'Believes' of ML_I into which it might be translated. Substitutions within the scope of the latter are less restricted. (Cf. note 12.)

$\ulcorner C_3(\text{Sense } (\ulcorner \text{Believes}_L \urcorner), a, b)\urcorner$ signifies the composition of the sense of 'Believes_L' with senses, *a* and *b*, of the expressions to which 'Believes_L' is applied. Such a composition is the proposition expressed by the sentence produced by predicating 'Believes_L' of its argument expressions.[13] We assume therefore these laws:

(A) $C_3(\text{Sense } (e_1), \text{Sense } (e_2), \text{Sense } (e_3)) = \text{Sense } (\ulcorner e_1(e_2, e_3)\urcorner)$

(B) $C_2(\text{Sense } (e_1), \text{Sense } (e_2)) = \text{Sense } (\ulcorner e_1(e_2)\urcorner)$

where '$\ulcorner e_1(e_2)\urcorner$' signifies the predicative application of e_1 to e_2. '$\ulcorner e_1(e_2, e_3)\urcorner$' is analogous.

By (a)–(d) it is easy to prove:

(i) $\text{Sat}(\alpha, \ulcorner \text{Believes}_L(\text{Igor}_L, \text{Believes}_L (\text{Bela}_L, \text{Masterpiece}_L (\text{Op. } 132_L)))\urcorner)$ $\leftrightarrow$ Believes (Igor, C_3(Sense ($\ulcorner \text{Believes}_L \urcorner$), Sense ($\ulcorner \text{Bela}_L \urcorner$), $A(\alpha, \ulcorner \text{Masterpiece}_L(\text{Op. } 132_L)\urcorner, O)$))

By (i), (e), (f), and (B), we have

(ii) $\text{Sat}(\alpha, \ulcorner \text{Believes}_L(\text{Igor}_L, \text{Believes}_L (\text{Bela}_L, \text{Masterpiece}_L(\text{Op. } 132_L)))\urcorner)$ $\leftrightarrow$ Believes (Igor, C_3(Sense ($\ulcorner \text{Believes}_L \urcorner$), Sense ($\ulcorner \text{Bela}_L \urcorner$), Sense ($\ulcorner \text{Masterpiece}_L (\text{Op. } 132_L)\urcorner$)))

It will be noticed that these proofs depend on allowing within the scope of 'Believes' (not 'Believes_L') substitutions of expressions that intuitively differ in sense.

Now from one point of view, neither (ii) nor the result of applying (A) to it can be our final explication in ML_I of the truth conditions of

[13] For heuristic purposes, I have taken ML_I to be first-order. As a consequence 'Sense (Believes_L)' does not occur in function sign position. This is contrary to Frege's would-be intentions. But the argument I am giving does not depend on this point. I shall assume that C_3 and C_2 are defined on the senses of syntactically appropriate expressions.

⌜$Believes_L$ ($Igor_L$, $Believes_L$ ($Bela_L$, $Masterpiece_L$(Op. 132_L)))⌝. (Call this sentence '(D)'.) For they would not explain the truth conditions of (D) to someone who understood ML_I but who did not already understand the expressions of L_I (did not understand what their senses are). Although (D) may perhaps describe the truth conditions, it does not 'give' them. To give the truth conditions, we need a sentence on the right side of the biconditional which states the relevant truth conditions purely in the terms of ML_I, without mentioning expressions of L_I. That is, we need a plausible translation of (D) into ML_I. From this viewpoint, an adequate theory of truth for L_I must satisfy Tarski's convention T, which requires that the metalanguage in which the theory is given provide a translation of the relevant sentence of L_I.[14] What might such a translation be?

It cannot be a sentence produced by translating component expressions of (D) by expressions in ML_I having the same denotation in their respective syntactic contexts. For suppose that we assume that the ordinary sense and denotation of '$Believes_L$' are the same as those of 'Believes'; those of '$Igor_L$' are the same as those of 'Igor', and so on. Then since L_I lacks a hierarchy, such a translation would yield a sentence amounting to (6). But an argument analogous to the one we applied to (6) earlier establishes that the purported translation is not equivalent to the original English sentence (5). So it would not be equivalent to any adequate representation of (5) in L_I.

Of course, we can reasonably question our assumption that the ordinary sense and denotation of '$Believes_L$' are the same as those of 'Believes' (and so on). But the grounds for questioning make reference to the same mismatches in structure between L_I and ML_I that we used to show that a term by term translation will not work. Giving up the assumption would make such a translation even less likely. It thus appears that there may be no way to 'give' the truth conditions of (D) in ML_I.

Let us summarize the difficulty. A language of the form of L_I cannot give a systematic truth theory for L_I—one needs laxer substitution principles than L_I countenances. Explications of truth conditions like (ii) do not satisfy the translation requirements of Tarski's Convention T. But term by term translations of embedded belief sentences of L_I into the 'laxer' metalanguage are prevented by the argument of section II. On the other hand, to attribute a hierarchy to L_I would be to controvert the hypothetical *informal* explication of embeddings in L_I, according to which no hierarchy is generated. Thus one might well feel that no such attribution would provide an intuitively acceptable *translation* of the relevant sentences of L_I. So Tarski's requirement is once again not met.

This is, I think, an interesting situation. It suggests a pair of options. The first is to emphasize the importance of Tarski's Convention T and hold that

[14] Alfred Tarski, 'The Concept of Truth in Formalized Languages', in *Logic, Semantics, Metamathematics* (Cambridge: Cambridge University Press, 1956), 187–188.

since (apparently) no informative and systematic truth theory for L_I meets it, L_I is an unlikely model of natural language since natural language should be capable (modulo the paradoxes) of interpreting itself. On this option, one would hold that an interpretation of a Method I language (as modeling natural language) would inevitably follow Frege's own lines and involve a hierarchy. For only thus will Tarski's requirement be met. A second option is to relax the requirement that an adequate theory of truth always provide translations (in some intuitive sense of good translation) of all the sentences of the object language. One might settle for semantical explications like (ii). These explain semantical structure without giving content. Assuming that one could master the content of the sentences of the object language by learning them directly, a theory of truth need, on this option, do no more.

A choice between the two methods must surely depend on global considerations governing the ultimate aims of one's semantical theory and one's representation of logical form. Insofar as lexical representation is our main concern, Method I has a strong claim on our attention. Insofar as we aim at a semantical theory which provides both a simple account of the formal structure of our language and a plausible representation of its content, Method II will remain dominant.

Postscript to "Frege and the Hierarchy" (2004)

Most philosophers who have thought about the matter take it to be unacceptable to be committed to a hierarchy of senses in accounting for embedded attributions of attitudes. Carnap rejected basic Fregean principles to avoid the hierarchy.[1] Dummett rejected what he claimed was an unimportant Fregean principle to avoid the hierarchy.[2] Early Terry Parsons maintained that Frege was not committed to the hierarchy, or at least could have avoided it without affecting the basic structure of his theory.[3] Davidson held that a language committed to a hierarchy is unlearnable.[4]

Church presented a detailed formalized language committed to the hierarchy. He believed, against Carnap, that substantial theoretical losses are incurred in avoiding it.[5] I believe that Church was right about this and that his point remains applicable to subsequent alternatives. I believe that the hierarchy is deeply grounded in Frege's standpoint, and that avoiding it requires giving up substantial parts of his theory. I also believe that avoiding it requires giving up principles that are motivated, powerful, and attractive—quite independently of Frege's maintaining them. Finally, I believe that it is not true that a language committed to an infinite hierarchy of senses is unlearnable.

I do not claim that there is no way around a sense hierarchy. I do believe that the reasons that favor postulating it have considerable plausibility. Senses are ways of thinking. So accounting for attribution of senses is accounting for attribution of (ultimately, thinking about) thought. I see Frege's account of

This Postscript benefited from two sessions in which I presented its ideas to the UCLA philosophy of language workshop in Winter 2004. Comments by Tony Anderson, Chris Peacocke, Jim Higginbotham, David Kaplan, Nathan Salmon, Philippe Schlenker, and Terry Parsons led to improvements.

[1] Rudolf Carnap, *Meaning and Necessity* (1947), (Chicago: University of Chicago Press, 1967), 131.

[2] Michael Dummett, *Frege: Philosophy of Language*, 2nd edn. (Cambridge, Mass.: Harvard University Press, 1981), 267 ff.

[3] Terence Parsons, "Frege's Hierarchies of Indirect Sense and the Paradox of Analysis", *Midwest Studies in Philosophy*, 6 (1981), 37–57. I say "Early Terry Parsons" to distinguish this view from later developments, cited in n. 20 and 24.

[4] Donald Davidson, "Theories of Meaning and Learnable Languages", repr. in *Inquiries Into Truth and Interpretation* (1965) 2nd edn. (New York: Oxford University Press, 2001). Strictly speaking, Davidson claimed that Church's language of the "Logic of Sense and Denotation" is unlearnable. I think that this point is doubtful, even applied to Church's particular language. (The issue depends on how Church's subscripts are related to a prior understanding of an upwardly-functionally-determined hierarchy of canonical concepts.) I believe that Davidson and many others have taken the argument to apply to any language that invokes a hierarchy of sense. Dummett does, for example, in *Frege: Philosophy of Language*, p. 167. What is clear, I think, is that the version of the hierarchy I outline here is not vulnerable to Davidson's objection.

[5] Alonzo Church, "A Formulation of the Logic of Sense and Denotation" (1951) in T. Burge *et al.* (eds.) *The Collected Works of Alonzo Church* (Cambridge, Mass.: MIT Press, forthcoming).

attribution of thought as a contribution to a scientific attempt to account for thought, including actual propositional attitudes, as well as abstract thinkable *Gedanken*. I believe that at present there is no superior account of thought, or the attribution of thought—insofar as that attribution fixes on the nature of thoughts and attitudes. So I believe that the "sense" hierarchy should be taken seriously—not merely as a historical curiosity.

Frege's notion of sense was *not* pointed primarily toward understanding linguistic meaning as a common denominator in a socially shared language. It was pointed at understanding thought expressed by language. Such thought is sometimes shared among language-users in a community, sometimes not. Frege was mainly concerned that it be shared in a *scientific enterprise*.

I do not accept all Frege's views about sense. I think that the contents of thought are not in general eternal and mind-independent, as Frege believed. There are other errors in his view.

I think that Frege was in many respects, however, on the right track. He was right in thinking that words express thought components that are abstract and that are distinct from the ordinary denotations or referents of the words or thought components. Thought components are representational, or have aboutness properties, and constitute an epistemic perspective on those referents. I also believe that Frege's appeal to oblique contexts in attributions of propositional attitudes is part of a correct view.[6] Although I will write here of Fregean senses, the structural points that I make about them are, in the main, applicable to abstract representational contents of thought. Most components of representational thought contents are what I call concepts.[7] Some

[6] In this paper, except where context makes another usage clear. I will use "oblique context" in such a way as not to pre-judge the issue between Method I and Method II formalizations of natural language. Method I formalizations take the same expression to express different senses and have different denotations in different contexts. Method II formalizations treat the same word-forms in natural-language identity and that-clause contexts as ambiguous, in such a way as to require formalization by different expressions. An oblique context O in this sense is a position in a natural-language sentence (construed in a certain way) in which intuitively substituting different word-forms that have the same denotation in certain other contexts (such as the contexts of singular terms in identity statements) is not guaranteed to preserve the truth-value of the whole sentence in which context O occurs. Thus, intuitively, substitution of "Samuel Clemens" for the second occurrence of "Mark Twain" in the sentence "Al believes that Mark Twain is identical with Mark Twain" is not guaranteed to preserve the truth of the whole sentence (on a certain construal of the sentence), although the two names have the same denotation in the identity sentence "Mark Twain is identical with Samuel Clemens". It is fundamental to Frege's method—and a view that I regard as fundamentally on the right track—that in oblique contexts a word-form denotes a sense, not the denotation that the same word-form has in contexts like those in an ordinary identity statement. I use "customary sense" similarly. It applies to the sense expressed by word-forms as they occur in non-oblique contexts. "First indirect sense" applies to the sense of an expression as it occurs in an unembedded oblique context. This usage leaves it open whether customary senses are identical with first indirect senses (and whether first indirect senses are identical with second indirect senses, and so on). According to the hierarchical view, all of these are distinct.

[7] This usage is, of course, distinct from Frege's use of "concept" (*Begriff*). My usage is closer to that of Alonzo Church, "A Formulation of the Logic of Sense and Denotation".

are what I call applications. Representational thought contents are true or false. Substituting "representational thought content" (used to include components as well as whole thought contents) for "sense" sometimes skirts common misunderstandings of Frege's view.

I will not discuss all these issues. I do want to say a little more than I did in "Frege and the Hierarchy" about how to think about the hierarchy in an intuitive manner. I also want to explain why a language committed to it is learnable. Finally, I want to discuss some more recent objections to the hierarchy and a response to my paper that attempts to avoid the hierarchy.

I will assume Frege's view that in ordinary, non-embedded attributions of propositional attitudes, expressions within natural-language that-clauses, occurring in positions where free substitution of co-denotational expressions is problematic, denote their customary senses, not their customary denotations. Consider again

(1) Bela believes that Opus 132 is a masterpiece.

Let us assume with Frege that the whole that-clause denotes the customary sense (thought content) of the words "Opus 132 is a masterpiece". The words within the that-clause denote their respective customary senses. These senses are components of the customary sense denoted by the that-clause.

It is evident that in using (1), one's understanding of the *customary* senses of the words in the that-clause plays a role in picking out those very customary senses. On Fregean principles, the customary sense cannot determine itself, since the customary sense determines a truth-value; and a given sense determines a unique s-denotation.[8]

Frege allows words in natural language to shift their senses and denotations with linguistic context. But he takes senses themselves to determine a unique entity (a unique s-denotation), if they determine any entity. He explicitly denies that the *sense* of an expression within the that-clause—its indirect sense—is the same as its customary sense.[9]

It seems to me that inasmuch as the entities denoted by words in the that-clause are different from those denoted by the same words in customary, direct contexts, it is intuitively evident that the modes of presentation associated with words in that-clauses are different from those associated with those words in direct contexts. This seems to me intuitively evident independently of appeal to Frege's principles. Thus "Opus 132" as it occurs in

[8] S-denotation or determination is the relation between a sense and the entity that it uniquely determines. That entity is denoted by an expression that expresses the sense. The relation of s-denotation or determination is the relation that Church called the *concept-of* relation. Determination of an s-denotation or a *determinatum* is the non-linguistic analog (between senses and entities rather than symbols and entities) of denotation. For Frege, and I think sometimes in fact, it is more basic than the relation that a linguistic expression bears to its denotation by way of its sense. Sometimes thought is more basic than language. Sometimes what is in fact a sense just happens to be expressed by a word. The thought content itself sometimes antedates expression through language.

[9] Gottlob Frege, "On Sense and Denotation", O 37.

customary contexts specifically denotes a string quartet. In a that-clause (occurring obliquely) it specifically denotes a way of thinking about, that string quartet. The sense, or way in which these different denoted entities are thought about should also differ. Similarly, "is a masterpiece" denotes a property of works of art, or (if you insist) a set of works of art, in customary contexts. In that-clauses, it denotes a way of thinking about that property. Ways of thinking about the property—modes by which thought is presented with the property—and ways of thinking about a way of thinking about the property are surely different.

One should stop thinking of senses as conventional meanings of words and sentences. Sometimes they are; sometimes they are not. Senses are ways of thinking, perspectives on entities presented, or purportedly presented, to the mind. The perspectives or ways of thinking about the denotations, or *determinata*, are different in the two cases. Given that these ways of thinking are not indexical ways of thinking, and given that the *determinata* are in the actual world, the difference in denotations or *determinata* argues very strongly for their being presented to the mind by different ways of thinking about them. To talk of a sense here as having multiple determinations or s-denotations is, I think, to lose sight of the fundamental role of sense in type-identifying specific cognitive perspectives on a subject matter.

I believe that "Frege and the Hierarchy" shows that if certain Fregean principles governing sense and denotation are maintained, then either in the object-language or in the metalanguage that explains the truth conditions of the object-language, one must distinguish the indirect sense of an expression (the sense it expresses in an unembedded that-clause) from its customary sense. I shall return to the details.

There does remain the fact that in using a that-clause one relies on one's understanding of the *customary* senses of the words within it in determining the that-clause's indirect denotation. Here the senses of the relevant words determine the representational content of Bela's thought. To understand the indirect sense of "Opus 132", one mainly has the customary sense (one's customary way of thinking about that string quartet) to go on. So the senses of the words as they occur in the unembedded that-clause must bear a relatively simple relation to the senses that they express when they occur in an embedded that-clause—the indirect senses.

One can see something similar going on in the way quotation marks are normally used. There is a legitimate but non-ordinary usage of quotation marks, especially common among philosophers of logic. On this logician's usage, quotations yield a name of the word-shape that they enclose, and one's understanding of the quotation abstracts completely from any meaning that the quoted word-shape might have. Of course, there are also uses of quotations of meaningless or foreign words that are not understood. But normally when quotation marks are used, there is an associated assumption that the quoted expressions have a meaning or sense in the language in which the quotation marks are used. This is why quotations are ordinarily translated.[10] For the sake of argument and illustration, let us include this assumption about the meaning or sense of the quoted word in what we shall call the normal construal or understanding of the quotation.

Consider this construal of the quotation expression " 'Opus 132 is a masterpiece' ". The expression denotes the expression "Opus 132 is a masterpiece"; and this expression is understood as having its usual meaning in English. The quotation expression " 'Opus 132 is a masterpiece' " does not have the same normal construal as the quoted expression "Opus 132 is a masterpiece". But one's understanding of the quotation expression depends only on understanding the shape and meaning of the quoted word, and on mastery of quotation marks. One's understanding of the quotation expression is determined by the sense of quotation marks, how many iterations of quotation marks are involved, and the sense and form of the words quoted.

That-clauses—at least insofar as they are used to characterize the *nature* of the propositional attitude—work like the normal construal of ordinary quotation expressions, just discussed, except that they are used to denote only the sense, or representational content, of words within them. Unlike quotation names, that-clause names do not denote the words *per se* at all.[11] Unlike quotation names, that-clauses enable one to denote just the sense shared by different words—for example, words in different languages. A non-embedded that-clause expresses a new sense that determines the customary sense. This new sense is the indirect sense of the word. This sense determines the customary sense of the words within the that-clause, whereas the sense of the words within the that-clause determine the customary denotations of the words. Normally these customary denotations are not senses at all. In fact, the logical type of a sentence standing alone differs from the logical type of the that-clause name. In Frege's view, sentences have truth-values as denotations. But, arguably, sentences do not denote anything at all: they simply have truth-values. An analogous point holds of some expressions within the sentence and the expressions that denote their customary senses: Their denotations may be of different type than the senses that they express. This

[10] For discussion of some of the complexities associated with this point, see my "Self-Reference and Translation," in M. Guenthner-Reutter and F. Guenthner (eds.), *Meaning and Translation* (London: Duckworth, 1978), 137–153.

[11] This view represents a change of position, on my part, from the position in "Self-Reference and Translation". I changed views relatively soon after writing that paper. At the time I wrote that paper, I was attempting to work out a Carnap-like position according to which words occurring obliquely in that-clauses denote words construed as having definite meanings. I have subsequently given up that view in favor of the view maintained by Frege and Church—that such words normally denote abstracta that are not necessarily linguistic entities, in fact, normally *not* linguistic entities. I do continue to believe that the rebuttal to Church's translation argument that I gave in "Self-Reference and Translation" both prevents the argument from being decisive, and indicates something important about the nature of quotation. The reason for maintaining the Church–Frege view about the denotation of words that occur obliquely is not Church's translation argument. It is that the Fregean view is more natural and provides a deeper account of what is really at issue in attributions of propositional attitudes—at least when those attributions are concerned with the natures of the attitudes being attributed.

is further reason to take the indirect senses to be distinct from the customary senses.

There is a function from the indirect sense of each expression to the customary sense of that expression: Sense determines denotation (or rather s-denotation, or *determinatum*), as usual. That is old news. The important thing to notice is this: The sense of the that-clause (understood to contain only oblique occurrences)—the first indirect sense of the sentence within the that-clause—is dependent on nothing other than the customary sense of the words within the that-clause, together with the sense of the that-clause-forming expression (expressed in English by "that") that is applied to those words.

We form a canonical name of the customary sense by applying the that-clause-forming expression ("that") to the sentence that follows. The word "that" is not crucial. The crucial things are the position of the expression in the subordinate clause and the intuitive point of the position to help specify a perspective or way of thinking. The indirect sense is uniquely fixed by this way of denoting the customary sense, assuming that the words in the that-clause occur obliquely. The sense of the canonical name can be uniquely recovered from the customary sense. There is a function from the that-clause name's denotation—the customary sense—to what I shall call the *canonical sense* that determines that customary sense. We understand the indirect sense, and "know what it is" in the sense of comprehending it, if we can both use the that-clause-forming expression ("that") and understand the sense of the sentential expression to which it applies. Nothing more is needed. I shall elaborate this point shortly, in the context of a formalization of the sense hierarchy.

So far we have arrived at two levels of senses—the customary sense and the indirect sense. Many philosophers have thought that one should stop with one level of sense, or at most these two. My paper shows, I think, that one cannot do so without giving up basic Fregean principles.

Given the way in which senses of these canonical names work, it seems to me evident that there is no particular problem about *learning* an infinite hierarchy of canonical senses expressed by such canonical names. The sense of a canonical name that denotes a sense at level n ($n \geq 1$) is a sense at level $n + 1$.

There is a *function* from a sense at level n to the *canonical* sense at level $n + 1$ that determines, or s-denotes, the sense at level n. Let us call this function "*the canonical sense* function".[12] Let us introduce the one-place

[12] This requirement can be modified. There may be reasons to allow for a more fine-grained conception of canonical sense and canonical name. Thus it has been claimed that word-forms with the same customary sense are substitutable in unembedded oblique contexts, but are not substitutable in embedded oblique contexts. Cf. Benson Mates, "Synonymity", in L. Linsky (ed.), *Semantics and the Philosophy of Language* (Urbana, Ill.: University of Illinois Press, 1952). To allow for such a view, one needs to allow that the indirect senses of word-forms with the same customary sense can differ; and word-forms that denote the same customary senses in oblique contexts can denote different

functional expression "C". The function denoted by "C" takes senses as arguments and yields, as values, *canonical senses* that determine or s-denote the argument sense. When "C" is syntactically applied to a *canonical name* of a sense s, it yields a canonical name of the canonical sense that determines that sense s.[13]

This functional expression could be roughly glossed "the canonical sense that determines", *except* that "C(s)" must be understood not as a description but as a canonical name.[14] "C" can, of course, be iterated.

Before discussing the sense of "C", let me say a bit more about the syntax of the language in which I conceive "C" as occurring. "C" will be used in formalizations of embedded oblique contexts in natural language. The that-clause-forming expression in English, "that" (like quotation marks), is not clearly a functional expression.[15] Nevertheless, the that-clause-forming expression "that" has many of the characteristics of a functional expression, especially in occurrences that embed other that-clauses. In a Method II formal language, one needs names for the customary senses of ordinary first-level expressions. Then a functional expression like "C" can syntactically apply to these names, and iterate.

indirect senses in singly embedded oblique contexts. I believe that this view can be accommodated by allowing that the indirect sense of a word-form is a function of the customary sense together with the word-form itself, or whatever is common to exact translations of the word which are finer-grained than sameness of customary sense. Then we would have a conception of a fine-grained canonical sense whereby understanding the fine-grained canonical sense that determines a customary sense requires understanding not only the customary sense, but what word expresses it. The conception of a fine-grained canonical name would parallel this notion of fine-grained canonical sense. Most of the examples used by Mates and his critics (including Church) seem to me to depend on conflating sense and conventional linguistic meaning. Thus *fortnight* and *period of fourteen days*, or *physician* and *doctor*, have respectively the same conventional linguistic meaning. But I think that they are commonly used to express different modes of presentation, different ways of thinking—hence different senses. Cf. my "Belief and Synonymy", *The Journal of Philosophy*, 75 (1978), 119–138. These matters need better sorting out. Nevertheless, it seems to me possible that there will be an explanatory use for the Mates-like position. I wish here just to indicate that the sense hierarchy can be conceived in such a way as to accommodate that position.

[13] I am comfortable in calling complex expressions formed in functional ways "names". They are not descriptions. They are rigid. I think that they are similar to numerals in the base-ten system larger than "9". But if one wants to require that names cannot have functional structure, or if these canonical expressions fail some other linguistic test for names, I am willing to call them "canonical designators". The key facts about them are that they are non-descriptive, rigid expressions, whose denotation can be computationally determined. There is a further feature of canonical names of senses that I am about to articulate in the *Principle for Canonical Names of Senses*. This feature or principle distinguishes them from other canonical names or designators, including numerals. My thinking about this matter goes back to "Self-Reference and Translation". Cf. note 10.

[14] Actually, the situation is slightly more complex. If "C" syntactically applies to any singular term T that denotes a sense, the resulting syntactical unit is a term that denotes the canonical sense. This is the canonical sense that determines the sense denoted by the singular term T. If "C" is syntactically applied to a canonical name of a sense, the resulting syntactical unit is a canonical name. If "C" is syntactically applied to any singular term other than a canonical name of a sense, the resulting syntactical unit is not a canonical name; and the sense of the whole functional expression is not itself a canonical sense.

[15] For quotation marks to be a functional expression, the expressions occurring inside quotation marks would have to be seen as naming, or otherwise denoting, themselves.

Suppose that the primitive expressions of the formal language initially do not include expressions that denote the senses of any of the language's expressions. Suppose that we add simple canonical names for the customary senses of each of the finitely many primitive expressions that are not themselves canonical names of senses. So we have doubled the number of primitive expressions. Then we add the functional expression "C". Thus we have canonical names for the customary senses of the primitive expressions, canonical names (via an initial functional application of "C") for the canonical senses that determine the customary senses, further canonical names (via one iteration of "C") for canonical senses that determine the canonical senses that determine the customary senses; and so on.

Later I will introduce a device for forming canonical names for the customary senses of syntactically complex expressions. Thus there will be means of composing canonical names for the customary senses of primitive expressions into canonical names for the customary senses of complex expressions. Further, there will be compositional principles for "C" that enable one to form canonical names for complex senses out of canonical names for the components of the complex. Thus, there will be principles for composing canonical names of complex senses to produce canonical names of the senses at any finite level of sense, from the first level, the level of customary sense, onward.

I turn now to the sense of "C". I maintain the following

Principle for Canonical Names of Senses: *The canonical name of a sense can be understood only if the sense that it names is understood.*

For example, to understand the sense of the canonical name for the customary sense of "3", one must understand the customary sense of "3".

I believe that the relevant canonical names obey a

Stronger Principle for Canonical Names of Senses: *To think the sense of a canonical name of a sense, one must simultaneously think the lowest-level (ultimately, customary) sense in the downward hierarchy associated with the canonical name.*

For example, to think the sense of the canonical name "C(C(s))", where "s" is a canonical name of a customary sense, one must think the customary sense denoted by "s".

These principles should be understood in light of the earlier discussion of quotation marks and of that-clauses. I regard expressions occurring obliquely in that-clauses as canonical names of senses. In unembedded that-clauses, such expressions are canonical names of customary senses. Customary senses are ways of thinking that determine or s-denote ordinary objects and properties (like string quartets, or violins, and properties of them). The customary sense or way of thinking that is denoted by an obliquely occurring expression in a that-clause is presented in a canonical way. The sense is thought about from the perspective expressed by a first-level canonical name of it. These canonical names have different senses from senses of other expressions that denote the same customary senses.

It helps here, as elsewhere, to remember that senses are fundamentally cognitive perspectives or cognitive modes of presentation. The canonical sense or way of thinking expressed by the first-level canonical name can be understood only by *simultaneously understanding* the sense or way of thinking that that expressed sense, or way of thinking, determines. (It determines the customary sense that the first-level canonical name denotes.) Thus to understand the sense of the first-level canonical name "Opus 132", as it occurs obliquely in an unembedded that-clause, one must simultaneously understand the denoted customary sense of "Opus 132". Similarly, to understand the sense of the first-level canonical name that *formalizes* "Opus 132", as "Opus 132" occurs obliquely in an unembedded that-clause, one must simultaneously understand the customary sense that the formalizing canonical name denotes.

The *Principle for Canonical Names of Senses* maintains in effect that understanding the canonical sense of any canonical name in the hierarchy obtained by iteration of "C", or of that-clauses, partly consists in understanding the customary sense of the expression at the bottom of the hierarchy. Here the analogy to the normal construal of quotation marks seems especially apt. To assume an understanding of the quoted expression, one must understand the expression at the bottom of the hierarchy of iterated quotations. Thus to translate """"masterpiece"""" into another language, one must understand "masterpiece".

It is clear that in attributing a way of thinking to someone in the that-clause fashion, one must understand that way of thinking. That is, in making reference to a way of thinking in an unembedded that-clause-type attribution, one must have a capacity to think with the way of thinking (the customary sense) that one attributes. One makes essential use of one's first-level mastery. But it is also clear that a capacity to attribute a way of thinking goes beyond the capacity merely to think with that way of thinking. It is one thing to be able to think a thought and another to be able to attribute the thought to someone else, or to oneself. There are stages of development when a child can do the first and not the second. Indirect senses expressed by canonical names of customary senses are ways of thinking or modes of presentation that mark or type-identify this additional attributive capacity that is parasitic on the capacity to engage in the root, first-level way of thinking. Higher-level ways of thinking expressed by higher-level canonical names make use of the same root, first-level capacity. They mark the additional difference, for example, between attributing an attribution and attributing a first-level way of thinking. It is one thing to attribute a thought and another to be able to think about someone else's, or one's own, attribution of a thought. Here again, different modes of presentation mark the different intellectual capacities. General theoretical principles, which I will discuss later, motivate distinguishing indefinitely many levels of potential perspectives, or attributional ways of thinking above these three levels. All such ways of thinking are canonical in that they are fixed (a) by what it is to understand and employ any attribution *at a given level*, and (b) by what it is to understand and employ the particular first-level ways of thinking that underlie the attributions.

There are some elementary things to re-emphasize about the sense hierarchy so far postulated. As with any sense of an expression, there is a downward function from a canonical sense to the denotation that it determines. Canonical senses determine unique senses. The canonical sense of any canonical name determines or s-denotes the denotation of the canonical name.

What is special to *canonical senses* is that there is an upward function from senses to *canonical* senses that determine them. Thus there is a "backward road" from senses to their canonical senses. The *Principle for Canonical Names of Senses* (and its stronger counterpart) should be helpful in getting an initial grip on this idea. Canonical senses are ways of thinking about senses that are grounded in a grasp of the senses that are thought about. I will be elaborating plausible principles that bring out that thinking about senses in certain contexts requires that the perspective on a sense—the way of thinking about it—be distinguished from the sense being thought about. This remains so even though the perspective is (largely) fixed by the subject matter, the sense, or way of thinking, thought about.

We began by introducing "C" as denoting a function from senses to canonical senses that determine them. Even granting that for each sense there is a unique canonical sense that determines it, there are many ways of thinking about this function. So this introduction does not in itself fix or explain the sense of "C". One might wonder which senses the canonical senses are. Or one might feel that although one understands the customary sense, it is mystifying what further canonical sense determines this customary sense.

I have gone beyond the introduction of "C" in terms of a description of the function that it denotes. I have indicated precisely what materials are used in understanding the higher-level senses. First, to understand the sense of a canonical name, one must know how to use names and have an understanding that distinguishes the sense of a canonical name from the sense of a description—even a description like "the sense of 'Opus 132' ". Second, I have invoked the *Principle for Canonical Names of Senses*. To understand the sense of a canonical name, one must also understand the customary sense at the bottom of the hierarchy within which the name is situated. Thus to understand the canonical sense that determines a customary sense, one must understand the customary sense. There is strong ground to distinguish the canonical sense that determines a customary sense from the customary sense itself. The customary sense is a way of thinking specifically about such things as string quartets or the property of being a masterpiece. The canonical sense that determines the customary sense is a way of thinking specifically about ways of thinking about string quartets or being a masterpiece. These ways of thinking–the ones about the quartets and such, and the ones about ways of thinking about quartets and such–are clearly different.

Now we go further in explicating the sense of the canonical names. To understand the initial layer of canonical names of the customary senses of primitive expressions in the language, one need only understand unembedded oblique occurrences of the expressions in natural language that these canonical names formalize. In understanding those natural-language expressions, one uses those expressions as canonical names (not descriptions). And one understands them by understanding the customary senses of those expressions (which are denoted by the expressions), while simultaneously understanding those expressions as denoting ways of thinking, not the entities (like string quartets) determined by the customary senses. Thus one is not using those customary senses as ways of thinking. One is using them as aspects of ways of understanding the ways of thinking that determine those customary senses. The full understanding of ways of thinking that determine those customary senses depends on comprehending attributions expressed (whether contingently or essentially) in unembedded oblique occurrences in natural language. Such understanding constitutes the additional layer of intellectual capacity discussed earlier. I believe that the senses of the initial layer of canonical names cannot be explained in any further way. No further way is needed.

To understand higher-level senses in the hierarchy, one must understand iterations of oblique contents signaled in English by iterations of that-clauses. Understanding "C" is essentially understanding the formation of a canonical name that meets the conditions of the preceding paragraph but also adds a level of perspective on the preceding two levels of senses. It is the understanding that accrues from understanding an iteration or embedding of a that-clause—from having the intellectual sophistication to take up the perspective of an additional layer of attribution. The key element is not the "that" itself. What matters is an understanding of canonical names informed by the *Principle for Canonical Names of Senses* and of the level of embedding of perspectives on perspective. "C" marks an embedding or a raising of canonical naming perspective. Iterations mark further levels of embedding. We understand "C" insofar as we understand the iterations of oblique occurrences in that-clauses that it helps formalize.

Thus I believe that it is necessary and sufficient in understanding canonical names formed with "C" that one understand (a) the customary sense, (b) a canonical naming perspective on that customary sense—a perspective of the sort involved in understanding sense-naming expressions occurring obliquely in unembedded natural-language that-clauses—and (c) a capacity to keep track of levels of iteration or embedding. Mastering the hierarchy of canonical names requires and involves nothing more. I believe that any other specification of the sense of "C", or of canonical names formed with it, is likely to be inaccurate or misleading. "C" mimics the intuitive understanding of natural-language embeddings of oblique contexts. If one understands that-clauses and their iteration, one understands the sense of such a functional expression, like "C", for building canonical sense-names.

What "C" brings out is that understanding higher-level senses is functionally dependent on (hence only on) four elements: (1) understanding the customary senses of the primitive non-sense-naming expressions; (2) understanding a finite number of canonical names of those customary senses (which requires only understanding a canonical name that "gives" the denoted sense by co-occurring with an understanding of the denoted sense);[16] (3) understanding (and keeping track) of levels of embeddings;[17] and (4) understanding principles for functionally composing canonical names of the senses of complex expressions from canonical names of the senses of simple expressions.

Thus it is necessary and sufficient to understand "C(C(<Opus 132>))"—where "<Opus 132>" is the canonical name of the customary sense of "Opus 132"—that one understand the larger expression and all its component singular expressions as canonical names, understand the underlying customary sense, and understand double embedding of an occurrence of "Opus 132".

Clearly, one can master this structure with finite resources: a finite number of canonical names for customary senses, functional applications of "C" to

16 Understanding a canonical name is to be distinguished from understanding a definite description. Unlike ordinary definite descriptions, canonical names are rigid designators. Unlike even rigidified descriptions like the "actual customary sense of "Opus 132"", they do not make reference to any general properties like "sense" or any entities other than senses (like words). These particular canonical names are special in two ways. One is that they involve situating the named entities in a structure, in something like the way that understanding the numerals involves situating the numbers in a structure. As with numerals, the canonical names that contain "C" are structurally complex, and the complexity matches aspects of the structure of entities that are named (by way of understanding the named sense). The other way that these canonical names are special is that they "give" the sense: understanding them requires understanding the sense that they name, and ultimately the root sense, in any given hierarchy. The key matter here is the *Principle for Canonical Names of Senses*. Cf. note 13.

17 Thus "C" just raises the level of perspective or embedding. "C" and its iterations are analogous to marking the number of digits in a base-ten numeral. The level of perspective has semantical import in itself, as does the number of digits. The difference is that the numerals do not relate to their denotations in as initimate ways as canonical names relate to *their* denotations (cf. the *Principle for Canonical Names of Senses*).

Here is a way that the level of embedding can matter semantically. In the sentence

(IS) Igor believes that Arnold's favorite proposition is something such that Bela believes it,

one can take "Arnold's favorite proposition" to occur obliquely. We reporters cannot substitute just any co-denoting expression for it in specifying Igor's beliefs. But from Igor's point of view, the term (or its sense counterpart) could occur transparently in Igor's thinking about Bela's beliefs. Arnold may not in fact have a favorite proposition, but Igor thinks he does. And Igor will allow substitution of any expression that he regards as co-denoting with "Arnold's favorite proposition" (or any sense that he regards as co-determining with the sense of "Arnold's favorite proposition") in his specification of Bela's belief. Thus our perspective must be distinguished from Igor's, and Igor's, in turn, from Bela's, even though the two perspectives which the expression "Arnold's favorite proposition" plays a role in specifying (Igor's and ours) understand the relevant canonical name of their respective senses fundamentally in terms of the root sense—the customary sense of "Arnold's favorite proposition".

those initial canonical names, and recursive principles for combining these canonical names into canonical names for the senses of syntactically complex expressions (like the senses of sentences).

There is no more difficulty in learning a language committed to an infinite hierarchy of senses than there is in learning a language that iterates quotation marks. There is no more difficulty in learning a language committed to an infinite hierarchy of senses than there is in learning a language involving base-ten canonical names for the numbers. The higher levels of canonical senses are determined by the customary sense together with iterations of the that-clause construction (or "C"). So all the talk by Davidson, Dummett, and others about our not knowing what the indirect senses and higher-level senses of expressions are, or of being unable to learn the hierarchy, is wayward and unsupportable.

According to Frege, each level of iteration of a that-clause (assuming that we are dealing with oblique contexts within that-clauses) yields a higher-level canonical name and expresses a higher-level canonical sense.[18] Thus

(2) Igor believes that Bela believes that Opus 132 is a masterpiece

attributes to Igor belief in the thought content that Bela believes that Opus 132 is a masterpiece. The that-clause that includes (1), as it occurs in (2), denotes the customary sense of (1). The customary sense of the expression (1) is composed of the customary sense of "Bela", the customary sense of "believes", and the customary sense of the that-clause "that Opus 132 is a masterpiece". The customary sense of this that-clause is a canonical sense that determines the customary sense of the sentence "Opus 132 is a masterpiece". So in (2) the that-clause "that Opus 132 is a masterpiece" denotes this level-2 canonical sense or cognitive perspective. It expresses a level-3 canonical sense that determines this level-2 canonical sense. And so on.

The higher-level senses denote ways of thinking about the lower-level senses. At each level one takes up a new attributional perspective on the level below. These ways of thinking, at different levels of canonical sense, differ, however, entirely in the level of embedding of attribution. Thus the differences are needed by, but are exhausted by, the levels of logical attribution. The fundamental "line" of content for each canonical upward route from a customary sense is fixed, *except for the levels of attribution*, by the customary sense itself.

Let me speak to two intuitive objections to the hierarchy. I believe that these objections are easily met, given the upward and downward functional relations among canonical senses.[19]

[18] Gottlob Frege, *Philosophical and Mathematical Correspondence*, ed. G. Gabriel *et al.* (Chicago: University of Chicago Press, 1980), 153–154; *Wissenschaftlicher Briefwechsel*, ed. G. Gabriel, F. Hermes, F. Kambartel, G. Thiel, G. Veraart (Hamburg: Felix Meiner, 1976), 234–237; Frege to Russell, 12/28/1902.

[19] Both of the objections that follow are given by Christopher Peacocke, "Entitlement, Self-Knowledge, and Conceptual Redeployment", *Proceedings of the Aristotelian Society*, 96 (1996),

Suppose with Frege that in (2) we make reference to a certain mode of presentation, partly fixed by a level of logical attribution of the belief content which Igor ascribes to Bela. Thus our attribution attributes to Igor a thought that contains a canonical sense that determines the customary sense of (1). The customary sense of (1) contains a canonical sense that determines the customary sense of "Opus 132 is a masterpiece". Assume that (2) is true. Thus Igor's belief contains a canonical sense that determines the (presumed) content of Bela's belief. And our attribution to Igor contains a (third-level) canonical sense that determines the second-level canonical sense that Igor (presumably) uses to think about Bela's (presumed) thought content, customarily expressed by "Opus 132 is a masterpiece".

One might object that this is an unintuitive description of the attribution in (2). One might object on the ground that for (2) to be true, (a) Igor must be thinking that one of Bela's beliefs has the content *Opus 132 is a masterpiece* and (b) in thinking this, Igor is employing the same content in thought as Bela would if he were to think that Opus 132 is a masterpiece. It might be further held (c) that Igor is not thinking of the content in some indirect way as the first content asserted on such and such a page of a particular, named book.

This objection is ineffectual. Frege would agree with (a). That is in effect just what (2) says. That is common ground. Frege would agree with (c). Canonical names do not denote "indirectly" as definite descriptions do. I believe that I have already developed this point in sufficient detail. Frege would agree with (b) *in a sense*. Igor could not attribute to Bela a belief through a canonical name of that belief content—or canonical sense that determines the belief content—unless he was also thinking the content. To attribute to Bela the content that he does, he must in doing so also think the content of Bela's presumed belief—the customary sense of "Opus 132 is a masterpiece". That is how canonical names of content work, as I have been explaining.

From a canonical perspective, we attribute to Igor a canonical perspective on Bela's presumed belief content. The perspective that we attribute to Igor determines that content. Our perspective determines the perspective that we attribute to Igor, and thereby indirectly determines that content. Further, Igor can have the perspective that we attribute only if he himself thinks the (attributed) content of Bela's thought. That is how he has a grip on his canonical concept of the content. Igor's attribution itself names the content. But naming it in this canonical way requires also thinking it. That is the burden of the *Stronger Principle for Canonical Names of Senses*. Understand-

142–144. Cf. also his *Being and Being Known* (Oxford: Oxford University Press, 1999), 245–62. He also makes an objection from an inference involving quantifying into that-clauses. I will not discuss this objection. I believe, however, that the hierarchical view is clearly at no disadvantage in dealing with quantifying in, as long as it is extended to allow for attributions of *de re* attitudes. Frege himself does not handle such cases. But the difficulties that his view faces do not derive from commitment to a hierarchy. I believe that his view needs to be supplemented to handle *de re* constructions and quantifying in. I believe that such supplementation need not essentially affect the structures that we are discussing here.

ing the canonical name requires understanding the customary sense that it names, as well as understanding the customary sense of the iterable construction for forming new canonical names. In our formalization, the iterable construction is "C". In natural language it is the first (or higher) iteration of the subordinate "that-clause" construction. This sense of "C" determines the function from senses to the senses that are canonical concepts of them.

In linguistic mode: Igor could not understand a canonical name for Bela's (presumed) belief content if he did not form the first-level canonical name from an exercised mastery of the expressions that customarily express that content. Here there is a close connection between that-clauses, quotation marks, and canonical sense-names. The advantage of the symbolic canonical names is that their iteration clearly maintains a maximum of extensional contexts. The advantage of the quotation marks and the that-clauses, as denoters of first-level, customary senses, is that they make it clear that understanding the naming device requires a prior understanding of the root named entity. It is a requirement on the symbolic canonical names (at the first level, hence at higher levels) that they be like that-clauses in this respect.

I think it clear that intuition cannot be expected to adjudicate whether the content of Bela's thought is contained in Igor's attribution or, on the contrary, determined by Igor's attribution, *where Igor's way of determining (thinking about) Bela's thought requires also thinking it*. Determination (s-denotation) is a theoretical concept answerable to structural as well as intuitive concerns. I believe that the Fregean view is equally in accord with intuition on the point. The Fregean view is strongly supported by the structural theoretical principles discussed in "Frege and the Hierarchy". It must be evaluated in terms of the power and explanatory value of those principles. (I will return to this issue.) I conclude that this intuitive objection is ineffectual.

A second intuitive objection goes as follows. From (2) and

(3) It is true that Opus 132 is a masterpiece

one can deduce:

(4) Something that Igor believes that Bela believes is true.

If one collapses the levels down to one, one can infer (4) from (2) and (3) by existential generalization. If one retains the levels, the objection goes, one can account for the deductive validity of the inference only through extremely complicated principles.

This objection is incorrect. By utilizing the downward functional structure of canonical names, once a formalization for sense-composition and canonical names is in place, it is a simple matter to produce a simple formalization of the inference. In fact, all that is needed is existential generalization. The key idea is that the embedded canonical (that-clause) name of the customary sense remains available to quantification because of the downward functional structure of canonical sense-names. Thus in the formalization of (2), Igor's attribution to Bela is specified with the functional expression: C(that Opus 132 is a masterpiece). One can quantify on the whole functional canonical name or on the argument place within it (or indeed onto the positions of the canonical names within the that-clause, which is composed of canonical names of senses).

To make these points a little more explicit: Let angle brackets yield canonical names of customary senses. Let the hat indicate appropriately formulated syntactic composition. Let (2) be formalized

(2a) Believes (I, <Bela> ^ < Believes> ^ C(<Opus 132 is a masterpiece>)).

Then by existential generalization and exportation:

(EG) ∃y(y = <Opus 132 is a masterpiece> & Believes (I, <Bela> ^ <Believes> ^ C(y))).

One can then use an obvious formalization of (3) to get (4). To capture the inference from (2) and (3) to (4), one need only quantify into the place of the canonical name "that Opus 132 is a masterpiece" (or "<Opus 132 is a masterpiece>") within the larger canonical name, "C(that Opus 132 is a masterpiece)" in (2a).[20] (For more on existential generalization applied to embedded constructions, see Appendix II.)

I have already explained the cognitive difference between the first-level customary senses and the (indirect) second-level canonical senses that determine them. Between the first and second levels, there are often differences in logical category, which clearly correspond to differences in modes of presentation. For example, the sense of a sentence is different from the sense of a singular term denoting that first-level sense. Moreover, each level of embedded attribution corresponds to different conceptual perspectives and intellectual capacities or levels of sophistication. The argument in "Frege and the Hierarchy" brings out why these distinctions of levels make a logical difference, not just in the relation between the first and second levels, but at each level of attributional embedding.

I believe that collapsing the hierarchy incurs serious costs in a formal representation of attributions of the representational perspectives that consti-

[20] A more nearly fully formalized language with these properties is set out by Terence Parsons, in "A Quasi-Fregean-Carnapian-Early Kaplanian Semantics", forthcoming in a volume honoring David Kaplan (Oxford: Oxford University Press). Parsons' up-arrow is essentially the same as my "C". We came to the idea independently. I agree with his remark that the hierarchy is "a kind of epiphenomenon of the simple part at the basis of the hierarchy"—with the proviso that one needs, in addition to an understanding of the simple part at the basis, a conception of a functional notion for yielding canonical senses (expressed by his up-arrow and my "C") to understand the cognitive content of the hierarchy. It is important that understanding canonical names at all levels depends on an antecedent *understanding* of the first-level customary senses. Parsons provides a detailed semantics in which he shows how to deal with inferences essentially similar to (2)–(4). He also discusses the argument of "Frege and the Hierarchy" and extends the argument in an illuminating way. I have some doubts about his account of variables relative to an assignment as being canonical names of the assigned objects. I believe, however, that this part of his view is not essential to formalizing relevant inferences.

tute propositional attitudes, and in understanding the relation between sense (or cognitive mode of presentation) and denotation.[21] I want to respond to one attempt to circumvent the argument of "Frege and the Hierarchy".

Christopher Peacocke tries to answer the argument I gave in "Frege and the Hierarchy" to show that Method I incurs serious costs if it is to escape the sense hierarchy.[22] Peacocke's view is that there is only one level of sense. He thinks that the senses of expressions in that-clauses are "redeployed". So the customary senses are not only denoted by obliquely occurring expressions in a that-clause. They are expressed by them as well. This is a version of Method I, one that seeks to avoid not only a hierarchy but even a second level of sense. It seeks to avoid treating indirect sense as distinct from customary sense. The only senses are customary, first-level senses. As indicated earlier, I think that when one reflects on what senses are—ways of thinking about purported referents—this view is one of the least plausible ways of avoiding the hierarchy. It seems to me far less plausible than the view that attempts to stop the hierarchy at two levels of sense. I am doubtful about either way, however.

My objection in "Frege and the Hierarchy" to taking Method I as a way of avoiding the hierarchy centered on difficulties that arise in giving a theory of truth for Method I.[23] Peacocke thinks that my objection rests on an oversight about how to give the truth conditions for natural-language belief contexts using Method I. He notes that although I say that the meta-linguistic semantical account that I gave for Method I does not "give" the senses of the object-

[21] Daniel R. Boisvert and Christopher M. Lubbers, in "Frege's Commitment to an Infinite Hierarchy of Senses", *Philosophical Papers*, 32 (2003), 31–64, offer an argument from Fregean principles for the hierarchy. Their paper does not make reference to mine. They center directly (and I think illuminatingly) on the functional-compositional structure of senses, whether denoted or expressed. They also make a more explicit textual case than I did that the principles that rule out a hierarchy can be found in Frege. Frege's commitment to the idea that one can recover the structure of thoughts from the structure of expressions expressing them occurs in various places. Cf. e.g. "Logic in Mathematics", *Posthumous Writings*, 207; *Nachgelassene Schriften*, 224. The authors do not make direct use of the substitution principles that I appealed to in an object-language (for Method II) or metalanguage (for Method I), as I did. And they do not discuss the Methods separately. But I explicitly associated the principles that I discuss with the functional-compositional structure of sense that they highlight. In one respect their argument is less general than my argument, or the extension of the argument that I will discuss. It does not apply directly to an attempt to collapse the hierarchy that appeals to only one level of sense. It shows on Fregean principles that if one is committed to two levels of sense, one is committed to the infinite hierarchy. They argue separately against a one-level theory. (I agree with some but not all of their objections to a one-level sense theory.) They also assume, appealing to Davidson's unlearnability argument, that Frege's commitment to the hierarchy of senses is untenable. As I have indicated, I believe that this assumption is very much mistaken.

[22] Cf. Peacocke, "Entitlement, Self-Knowledge, and Conceptual Redeployment", 153–157.

[23] I continue to think that this objection to Method I raises interesting questions about the relation between a truth theory for a language and the structure of that language. I continue to think that the objection is correct. But I believe that there is an objection to taking Method I as a way of avoiding the hierarchy that is more direct. Such an objection works directly off the alleged senses and denotations of the parts, without using any "syntactical" premise about substitutivity. The argument that I gave in the original paper, using a premise about substitutivity in contexts that correspond to oblique contexts in attributions of propositional attitudes, was envisioned as applicable to the metalanguage that gives the truth theory for the Method I. But the last argument I give below can be adapted to apply directly to

language, but only describes them, this should not be surprising. For the truth-theoretic axioms that I give (in Section III of "Frege and the Hierarchy") do not say what the senses of the expressions of the language are. He then goes on to give a truth theory that does give or specify the senses.

This answer rests on a misunderstanding of my objection. I never doubted that one can give or specify the senses of the object-language and carry through a truth theory using those specifications. In fact, under the rubric of "translation" into the metalanguage I specifically discuss a truth theory that specifies senses. My claim (in section III of "Frege and the Hierarchy") was that in a metalanguage that allows substitutivity of co-denoting expressions and that specifies senses by translating expressions of the object-language (in particular ones in oblique contexts that specify senses) into metalinguistic expressions with the same senses, the problems I raised for Method II will recur in the metalanguage. They will recur in a metalanguage capable of providing a systematic semantics for the truth conditions of sentences of a Method I object-language. The problems will recur if one tries to avoid the hierarchy, unless one gives up basic plausible principles about sense and denotation.

The main relevant assumptions invoked in the original article are of three types. One is that the extensional substitution principles needed in a truth theory are maintained. This is the syntactical counterpart of the principle that a sentence's truth-value is fixed by the denotations of its parts. The syntactical counterpart is needed to carry out standard proofs within a truth theory for any ordinary object-language. A second type of principle is that sense determines denotation. For the argument, I needed only an instance of this principle: that sentences with the same sense have the same truth-value. The third type is that the translations are compositional and that principles of functional composition and decomposition of senses are maintained in the metalan-

an object-language that follows Method I, without making use of any premise about substitutivity that Method I bars. Method I directly incurs a hierarchy on plausible principles. I outline this more direct argument in *Appendix I: Direct Pressure on Method I.* (It is probably advisable to read the Appendix after reading through the main text.)

Contrary to many who work on these issues, I do not think it at all obvious that Method I is a closer and more natural formalization of natural language than Method II. The issue is how to account for the special sort of "ambiguity" that arises between non-oblique contexts and oblique contexts. Most formalizations of ambiguous terms introduce different terms that correspond to the different "meanings" of a single word-form—just as Method II does. It is true that the type of ambiguity involved in oblique contexts is structural and systematic. It is also true that in oblique contexts, where the customary sense of a word is denoted, the denoted sense plays an essential role in understanding the sense of the occurrence that denotes the customary sense. So the different senses of a word-form in natural language, which are expressed in oblique and non-oblique contexts, are bound together. But Method II as well as Method I incorporates into its account an acknowledgment of this point. In Method II, the acknowledgment lies in the way that understanding higher levels of sense depends on understanding the first level. The issue over the best account of *natural language* ultimately depends on which formalization best accounts for inferences. I venture to predict that Method II will be found to be superior on this score. What I believe with more confidence is this: Method II languages are superior as frameworks for scientific and philosophical theories of thought.

guage. Peacocke does not investigate my claim that the hierarchy will recur in the metalanguage that gives the truth theory for a Method I object-language, unless one gives up one of the principles. (I held that the substitution principle cannot be given up if one is to prove the theorems of the truth theory.) In particular, he does not investigate whether the hierarchy will recur in his own metalanguage. Thus he does not even address the claim of section III of the 1979 paper, much less answer it.

I will sketch some arguments that show how the metalanguage that Peacocke relies upon threatens to lead to the same hierarchical result that I discussed with respect to Method II languages. The first argument that I will give faces a pair of prima facie problems. I will show that the two problems are not fundamental by considering two further arguments. The last of these arguments seems to me to show decisively that my original criticism of the position applies to Peacocke's metalanguage. A fully formalized version of the arguments would be more perspicuous, but would also take up more space. I will provide enough formalization to make them clear.[24]

Let "$<\ldots>$" be a canonical name of the customary sense of the expression that fills in for the dots. (This is Peacocke's terminology.) So "$<\text{Opus132}_L>$" denotes the customary sense of "Opus 132" as it occurs in the Method I, formalizing object-language L. Let "$<\ldots> {}^\wedge <\ldots.> {}^\wedge <\ldots..>$" denote the result of functionally composing the senses named by the bracketed expressions into a canonical name of the grammatically appropriate complex expression.[25] "$<\text{Opus132}_L> {}^\wedge <\text{Masterpiece}_L>$", for example, canonically names the customary sense of the natural-language sentence "Opus 132 is a masterpiece", as those expressions (subscripted by "L") occur in the Method I, formalizing object-language L. Such compositional principles apply to the translations of L's

[24] I give two semi-formal arguments here. The first argument is an application to Peacocke's language of a simplified variant on my 1979 argument. The second is an application of an extension of my 1979 argument by Parsons, "A Quasi-Fregean-Carnapian-Early Kaplanian Semantics", to singular terms. I think that the argument generalizes further. My original argument focused on sentences within that-clauses. But the basic ideas are the same, as applied to any non-sentential expressions that have sense, not just sentences and singular terms.

An anticipation of the arguments of the present Postscript, pointed out by Tony Anderson, can be found in Leonard Linsky, *Referring* (New York, Humanities Press, 1967), 35. Linsky uses the principle that sentences with the same sense have the same truth-value, as do I. But his other premise is different from the two I use (compositionality and decomposability of sense). For both present dialectical purposes and purposes of clarifying fundamental principles about specification and explanation of thoughts, I prefer my argument. But Linsky's argument is of independent interest.

[25] The notation "$^\wedge$" reads "appropriately grammatically composed with". "$^\wedge$" is a wave of the hand toward what will inevitably be an extremely complex account. The account must specify the various ways in which one puts together names of complex senses of complex expressions, where the complex expressions have components of different syntactical categories. All that I assume here is that canonical names can be formed in such a way that the sense of the name of the sense of a complex expression is a function of the senses of the expressions that are components of the complex expression. The name-forming rules will make use of the syntactical rules governing the expressions with the senses named, and the ordering of the syntactical parts (with their senses) within the syntactical complex.

sentences into the metalanguage. So for example "$<\text{Opus132}_L>{}^\wedge <\text{Masterpiece}_L>$" also names the sense of the metalanguage sentence "Masterpiece(Opus 132)", since this metalanguage sentence translates the object-language "Opus 132_L is a masterpiece_L". And "$<\text{Masterpiece(Opus132)}>$" names this same sense, by way of the metalinguistic expressions that express it.

According to the view that I oppose,

(a) The customary sense of "Bela believes that Opus 132 is a masterpiece" $= <\text{Bela}_L>{}^\wedge <\text{Believes}_L>{}^\wedge <\text{Opus } 132_L \text{ is a masterpiece}_L>$.

I assume (a) for *reductio*. (a) combines the beginning of a functional decomposition of the sense of (1), "Bela believes that Opus 132 is a masterpiece", with the view that in the that-clause of (1), the *customary* sense of "Opus 132 is a masterpiece" is expressed. That is, what is expressed is the customary sense, rather than a further indirect sense.[26]

Now suppose that the sense of "Opus 132 is a masterpiece" (in the natural language) and the sense of its counterpart in the Method I, formalizing object-language are the same as the sense of the sentence that translates it, "Masterpiece(Opus 132)", into the metalanguage. And suppose that the senses of "Bela_L" and "Believes_L" are the same as their metalinguistic counterparts. So translation is normal but tries to capture Peacocke's redeployment view of the object-language L, according to which there is only one level of sense.

Then

(b) $<\text{Bela}_L>{}^\wedge <\text{Believes}_L>{}^\wedge <\text{Opus } 132_L \text{ is a masterpiece}_L> =$
$<\text{Bela}>{}^\wedge <\text{Believes}>{}^\wedge <\text{Masterpiece(Opus 132)}>$.

(c) $<\text{Bela}>{}^\wedge <\text{Believes}>{}^\wedge <\text{Masterpiece(Opus 132)}> =$
$<\text{Believes(Bela, Masterpiece(Opus 132))}>$.

(c) follows from (b) by the functional compositionality of sense. The sense of the expression within the angle brackets is functionally dependent on the senses expressed by the constituent parts of the expression.

(d) The customary sense of "Bela believes that Opus 132 is a masterpiece" $=$
$<\text{Believes(Bela, Masterpiece(Opus 132))}>$.

(d) follows from (a)–(c) by transitivity of identity.

[26] Peacocke's account assumes that the that-clause denotes as well as expresses its customary sense. I have mentioned earlier that I believe that this is an unattractive feature of the account—certainly out of keeping with the Fregean view that a sense determines a unique denotation. On his view the sense is a mode of presentation both of a sense and of a truth-value. Or if one rejects truth-values, the sense associated with "Opus 132" is a mode of presentation both of a sense and of a string quartet. Such a view seems to me to lose any plausible connection to the idea that senses are modes of presentation or ways of thinking. This view also loses connection between senses (as thought components) and type-indentification of specific cognitive abilities. The ability to think about a string quartet is different from an ability to think about a sense.

(e) "Bela believes that Opus 132 is a masterpiece" is true in L $\longleftrightarrow$ Believes(Bela, Masterpiece(Opus 132)).

(e) follows from (d) by the principle that free-standing sentences with the same sense are materially equivalent. Suppose that "Masterpiece(Opus 132)" is true. Let "S" be any other true sentence in the metalanguage. Then

(f) "Bela believes that Opus 132 is a masterpiece" is true in L $\longleftrightarrow$ Believes(Bela, S).

(f) follows from (e) by the extensionality of substitution principles in the metalanguage and substitution of "S" for "Masterpiece(Opus 132)".

Similarly, since "Opus 132" occurs in extensional position in the metalanguage sentence that expresses the sense <Bela>^<Believes> ^ <Masterpiece(Opus 132)>, singular terms with the same denotation are interchangeable in sentences that express attributions of belief.

(f) is a *reductio ad absurdum*. What does it reduce to absurdity? The principles of the functional compositionality and decomposability of sense and the principle that if the senses of sentences are the same, they have the same truth-value, seem very plausible. The extensionality of substitution principles in the metalanguage seems necessary for giving a systematic theory of truth.

I see three ways of pinning the absurdity of the conclusion on something other than the assumption that the customary sense of a sentence is expressed in oblique occurrences in object-language that-clauses. I believe that none of these ways will ultimately be satisfactory.

One way is to claim that it is *ungrammatical* to attach "Believes" syntactically to a singular term ("Bela") and a sentence ("Masterpiece(Opus 132)").[27] Such attachment constitutes putting together senses of expressions in a compositional way. In fact, it composes the very senses that Peacocke claims are involved in the belief attribution. It matches the claimed sense of the object-language's that-clause with a sentence in the metalanguage with the same sense. But one could maintain that the compositionality principle is inapplicable at step (c), because the resulting string is ungrammatical. I believe that this claim will not suffice to escape the difficulties, and that it does not go to the heart of the matter.

Frege took sentences to have the same sense and denotation as certain terms—nominalizations of the sentences. We can easily recast the argument so that we use only nominalizations of sentences in that-clause position. We maintain the principle that sentence nominalizations that have the same sense

[27] There is, in my view, an unclarity about the logical syntax of "believes that *p*" in the object-language if one accepts the view that the sentence in the that-clause both denotes its customary sense and expresses its customary sense. Insofar as it expresses its customary sense, it is a sentence. Insofar as it denotes its customary sense, it is a term. I believe that this unclarity is an aspect of the fundamental problem. I believe that in actual fact the that-clause is a singular term, and "believes" is a relational predicate.

denote the same truth-value—and then carry through the same argument. It seems clear that this doctrine of Frege's is not the key issue in whether or not there is a hierarchy.[28]

Moreover, an argument exactly parallel to (a)–(f) can be applied to the object-language sentence "Byron searched for Ossian". In direct contexts, "Ossian" and "Tlaloc" lack a denotation. "Byron searched for Ossian" can be true while "Byron searched for Tlaloc" is false. Other singular terms can be used to make the point without using ordinary proper names.

The Fregean treatment holds that a term that follows "searched for" that resists coextensional substitution denotes its customary sense. Here there is no change from the grammar of direct contexts to the grammar of indirect contexts. There is a singular term in both cases.

Now consider iterations such as "Browning questioned Byron's search for Ossian" and "Eliot questioned Browning's questioning of Byron's search for Ossian". The analog of the original argument still produces an absurd conclusion ("Bela searched for Ossian" is true in L if and only if Bela searched for T, where "T" is any term coextensive with "Ossian", or coextensive with whatever other term occurs in the position of "Ossian"). This is an absurd consequence that does not at all depend on the ambivalent grammatical position that the view under discussion involves (and forces) in its explanation of the roles of expressions in ordinary that-clauses. So the grammatical issue cannot be fundamental.

So I take the attempt to block (a)–(f) by appeal to the ungrammaticality of the key sentence not to lead anywhere worthwhile.

A second way to attempt to block the argument on behalf of the one-level theory is to allow that there are two relevant "belief" sentences in the metalanguage. One is the one that we have been discussing—"Believes(Bela, Masterpiece(Opus 132))". The other is the same except that instead of a sentence in the second argument place of "Believes", there is a singular expression. This singular expression must both denote and express the customary sense of the sentence, "Masterpiece(Opus 132))". Thus, in the metalanguage, there is a singular term that denotes the customary sense of the sentence, but expresses that same sense. Then one simply denies that "Believes (Bela, Masterpiece(Opus 132))" is true. Only the belief sentence with a singular term in the second argument place can be true. This view tries to block the argument at step (e). It rejects the principle that free-standing, nonindexical grammatical sentences with the same sense have the same truth-

[28] In "Frege on Truth" (1986) (Ch. 3 above) I maintain that the usual reasons for holding that Frege was mistaken in giving sentences and singular terms a single grammatical category *for some purposes*, are not good ones. I do not think that there are deep philosophical reasons against Frege's view on this point or against his view that sentences and their nominalizations have the same sense. Still, I do not accept either of these views. So I do not ultimately rest my case here on siding with Frege against this complaint. Nevertheless, I think that the complaint is insubstantial and does not go to the heart of the matter.

values. For "Believes (Bela, Masterpiece(Opus 132))" and "Believes (Bela, ST)" have the same sense but different truth-values (where "ST" stands in for a singular term that denotes the sense of "Masterpiece(Opus 132)" and expresses that same sense). I believe that this way of attempting to block the argument is unacceptable. The principle that is rejected is a very basic and plausible one. In fact, this way of attempting to block the argument involves commitment to quite a number of further unattractive consequences. I will return to these and discuss them in the context of a further argument ((a′)–(f′), below).

A third way of attempting to pin the absurdity on something other than the collapse of the hierarchy is to reject an assumption behind step (b) of the argument—the assumption that "Believes" in the metalanguage has the same sense as the object-language "Believes_L". One can plausibly argue as follows. It turns out that, in the metalanguage, "Believes" works so that if a person "Believes" (in scare quotes) one truth, he or she believes them all. So what could show more clearly that the metalanguage expression "Believes" does not have the same sense as the object-language "believes" ("Believes_L")? Similarly, the absurd results in (e) and (f) simply bring out that the truth theory rests on bad translation from the object-language into the metalanguage. Something is peculiar about either the supposed senses of the sentences or the attempt to match them in translation.

Now I think that this point is correct. But the question is what is leading to this state of affairs? We used translation of the sentence "Opus 132 is a masterpiece" into "Masterpiece(Opus 132)", which seems entirely correct. Those sentences in the respective languages (object-language and metalanguage) express the same sense. And we applied the composition principle with the resources that we had. We need to consider whether good translation is possible, consistent with the principles that we are assuming. This was the dialectical situation that I envisioned in "Frege and the Hierarchy". The argument developed so far is a minor variant on the one given in the original 1979 article. I claimed that if one produces a translation, one will either find the hierarchy in the metalanguage, or run afoul of the relevant Fregean (and otherwise plausible) principles. Let us investigate this matter further. I believe that the problem in translation is just a symptom of the more fundamental disease.

What we need in the metalanguage is a singular term that denotes the customary sense of "Masterpiece(Opus 132)", but does not introduce a further indirect sense. It expresses the sense that it denotes. Perhaps this is the analog in the metalanguage of the way the object-language's that-clauses are supposed to work. This is an idea that we came upon in discussing the first possible difficulty with (a)–(f), the grammatical difficulty. So let us follow it out here.

We continue to take "<Masterpiece(Opus 132)>" to be a singular term in the meta-language that denotes the customary sense of the sentence "Masterpiece(Opus 132)". This singular term is now to be taken as having no further sense: its sense is the same as the sense of the sentence.

This is in itself an absurdity. The sense of the sentence cannot be the same as the sense of a singular term that is about a sense or a way of thinking. The sense of the sentence is a way of thinking about a string quartet and about its being a masterpiece. The sense has components that determine a string quartet and its being a masterpiece. The sense of the relevant singular term, "<Masterpeice(Opus 132)>", is a way of thinking about these ways of thinking.

One needs to bear firmly in mind what senses are. They are perspectival ways in which a referent (if any) is presented to the mind. The problem is not merely that the one-level-of-sense view gives up the principle that a sense determines a referent or denotation. It is that by the nature of sense, a (non-indexical) sense cannot be a single way of presenting such different denotations to a mind.[29] Clearly we are talking about different modes of presentation of different denotations. The idea that the sense of the term that denotes the customary sense of "Masterpiece(Opus 132)" does not introduce a further sense beyond that customary sense of the sentence is, I think, unacceptable. But I will follow out the idea in the context of our principles to show where it leads.

Peacocke does not discuss the senses of the canonical sense-names that he introduces into the metalanguage. But it is obvious that on Fregean theory, they have to have senses. They express different ways of thinking about senses than, say, definite descriptions or other names of the senses that we might whimsically introduce. We should be able to form canonical names of the senses expressed by these first-level canonical names. Thus it would seem that we can regard the expression "C(<Masterpiece(Opus 132)>)" as a canonical name denoting the canonical sense of the metalinguistic expression "<Masterpiece(Opus 132)>". On the one-level anti-hierarchy view, the sense denoted by the first expression will be the same as the sense denoted by the second. (Or perhaps, one would just bar introduction of the first expression.) Of course, we can also produce informal specifications of the sense of "<Masterpiece(Opus 132)>", as follows: the sense of

[29] In my own view, certain ways of thinking are token applications, or abstractions from token applications. These do not determine their denotations by their nature. They depend on context for their successes. But for any given token application there is at most one denotation or referent. So even these "ways" of thinking are modes of presentation of an object (if any) that are individuated in such a way that a given way of thinking determines at most one denotation. Here way of thinking determines a unique referent, if any, in the functional sense of "determines". This is, I think, the sense of "determines" that is fundamental for understanding Frege and for understanding representation. For more on this, see my "Belief *De Re*", *The Journal of Philosophy*, 74 (1977), 338–362; "Russell's Problem and Intentional Identity" in Tomberlin (ed.), *Agent, Language, and the Structure of the World* (Indianapolis: Hackett, 1983), 79–110; "Vision and Intentional Content" in E. Lepore and R. V. Gulick (eds.), *John Searle and his Critics* (Cambridge, Mass.: Blackwell, 1991); and "Five Theses on *De Re* States and Attitudes", in a forthcoming volume in honor of David Kaplan (Oxford: Oxford University Press).

"<Masterpiece(Opus 132)>". But in my view, these specifications do not have the same sense as the canonical names.

As I envisage a fully formal language, I postulate an initial layer of canonical names for the sense of each primitive expression of the Method II language. Then there will be canonical names, formed with "C", for the sense of those canonical names, and iterations for higher-level senses of the sense of the expression formed with "C", and so on. Thus, "C(<Opus 132>)" is a canonical name of the canonical sense that determines the sense denoted by "<Opus 132>". (Of course, "Opus 132" does not occur as a name in extensional position within the angle brackets.) There are also composition principles for denoting the sense of a complex expression on the basis of the senses of its parts.

If it is to avoid, in the truth-theoretic metalanguage, exactly the same hierarchy of senses that Frege makes use of to account for iterated contexts, the single-level view must claim that for canonical names of senses, the senses of the names are identical with the senses that they denote. This view captures in the metalanguage the way that expressions in that-clauses operate in the object-language—according to the view that collapses the hierarchy to one level (the redeployment view): The expressions both denote and express the expressions' customary senses.

Such a view may seem to make only a small concession to the broadly Russellian idea that names have denotation but no (new) sense—or that sense is identical with denotation. The concession is only for special names, canonical names of senses. For these names, the sense and denotation are the same. The concession appears to be fairly close to Russell's restricted view of "logically proper names," which supposedly name universals that we "grasp" by acquaintance. In fact, this seemingly small concession leads to very serious trouble.

Here is a further argument that is relevant to understanding the difficulties I raised for a metalanguage that specifies senses (or translates object-language expressions into the metalanguage), gives a truth theory for Method I object-languages, and attempts to avoid the hierarchy. The argument is that either the metalanguage is committed to a hierarchy of senses after all, or it falls again into collapsing sense and denotation.

For simplicity, I will carry out the argument on the sense named by "<Opus 132>", rather than on the sense named by "<Masterpiece(Opus 132)>". This will circumvent the grammatical issue discussed earlier. I believe myself to have shown that that issue is not fundamental. But just to keep matters simple, I will avoid it in this argument from the beginning.

In what follows I will use quotation marks rather than canonical names for the expressions of the Method II language. In a fully formal exposition, I would avoid this. But for present purposes, I believe that quotation makes for easier reading. The argument could easily eliminate quotation marks without any substantial change.

Suppose, according to the view that collapses the hierarchy and that makes the concession to Russell:

(a′) The sense of "<Opus 132>" = <Opus 132>.

(a′) says that the sense of the canonical name, "<Opus 132>", which denotes the customary sense of "Opus 132", is not an additional sense. It is the customary sense "redeployed". According to (a′), the sense of the term that translates an unembedded obliquely occurring expression is identical with the (customary) sense denoted by an unembedded obliquely occurring expression. (a′) makes "<Opus 132>" match in the Method II truth-theoretic metalanguage the semantical behavior of the counterpart expression which it translates from the Method I language, or from the obliquely occurring expression in natural language.

As noted earlier, I think that step (a′) is already absurd. A way of thinking about a string quartet cannot be the same as a way of thinking about a way of thinking. Certainly no theory of thought should be committed to this. Methods of individuating cognitive abilities attributed by a scientific or philosophical theory of thought need to be more flexible and fine-grained. In cases of successful representation, they need to individuate ways of thinking in terms of what they are about.

This step, (a′), can be motivated by asking the proponent of a one-level theory of sense to choose the expression in the truth-theoretic Method II metalanguage that most closely translates expressions that formalize in the Method I language the obliquely occurring expressions in the natural language (or the expression in the Method II metalanguage that most closely translates English obliquely occurring expressions).[30] "<Opus 132>" is the term in Peacocke's truth-theoretic, Method II metalanguage that he uses to translate the obliquely occurring expression "Opus 132" in English (or in the counterpart expression in the Method I object-language). I do not know of a better choice. It is a canonical name that "gives" the sense. It does not merely describe it. One must understand the sense denoted (the customary sense) if one is to understand this name of the customary sense. In these respects, the canonical name "<Opus 132>" matches the English obliquely occurring expression and its Method I counterpart expression. Given that this canonical name is the closest translator in the Method II language, and prima facie a good translator, we can ask what the sense of "<Opus 132>" is. According to the view under discussion, the sense of the obliquely occurring expression

[30] One of the possible "ways out" from the *reductio* about to be developed is to claim that English and a Method I language do not translate into a Method II language. On such a view, no truth-theoretic semantics that relies on translation can be given for English or for a Method II language. One could hold that one can give a semantics for English or Method I, but no translational semantics (i.e. no truth theory that respects Tarski's schema where the right side translates the object-language sentence mentioned on the left side). The *reductio* proof that I am about to give does not purport to defeat such a view directly. But I do not think denial of translation is plausible. At any rate, Peacocke does not take this "way out".

in English is identical with its oblique reference, and there is no need for a further sense. To stay as close to the way that the term that it translates is supposed to work, we say, in (a′), that the sense of "<Opus 132>" is nothing other than the customary sense of "Opus 132".

Then,

(b′) The sense of "<Opus 132> = <Opus 132>" = the sense of "<Opus 132>" ^ the sense of "=" ^ the sense of "<Opus 132>".

(b′) depends only on the principle that the sense of a whole sentence (the trivial identity sentence) can be functionally decomposed into the senses of its parts.

(c′) The sense of "<Opus 132>" ^ the sense of "=" ^ the sense of "<Opus 132>" = the sense of "<Opus 132>" ^ the sense of "=" ^ <Opus 132>.

(c′) follows from (a′) and (b′) by substitutivity of identity (here applied in a context that does not translate or otherwise formalize a natural-language oblique context).[31]

(d′) The sense of "<Opus 132>" ^ the sense of "=" ^ <Opus 132> = the sense of "<Opus 132> = Opus 132".

(d′) follows from (c′) by functional compositionality of senses. The sense (or sense-proposition or thought content) composed appropriately of the senses of the semantically relevant parts of a sentence is identical with the sense (or sense-proposition or thought content) expressed by the whole sentence.

(e′) The sense of "<Opus 132> = <Opus 132>" = the sense of "<Opus 132> = Opus 132".

(e′) follows from (b′)–(d′) by transitivity of identity.

(f′) <Opus 132> = Opus 132.

(f′) follows from (e′) by the truth of self-identities, propositional calculus, and the principle that free-standing sentences with the same sense have the same truth-value.

The argument is replicable at any higher level of the hierarchy. So it applies to any view that identifies doubly indirect senses and indirect senses.

[31] A variant of the argument that I am giving is applicable directly to Method I; see Appendix I. The style of argument that I am giving is not restricted to application to a Method II metalanguage. But here I am illustrating a variant of the argument that I gave in "Frege and the Hierarchy" regarding a Method II translational truth theory for English or for a Method I language.

It should be emphasized that we need not have used a canonical name like "<Opus 132>" that makes use of an "internal" occurrence of the name for the customary referent. Any canonical name of the sense of the name for the customary referent would do, as long as it respects the principles governing canonical names for senses.

Analogs apply for other expressions besides names and singular terms. So an analog will apply to senses of predicate expressions.

So if one makes what initially may have seemed to be a small concession to a Russellian conception of canonical sense-names; and if one maintains plausible general principles about sense (especially plausible if one thinks of sense—as one should—as way of thinking), the distinction between sense and denotation for singular terms, indeed for all expressions, collapses across the board. The argument assumes only the functional compositionality and decomposability of senses, substitutivity of identity in contexts that do not translate natural-language oblique contexts, and the principle that the senses of sentences determine their truth-values. This last principle need be applied only to non-indexical sentences standing alone and lacking any analog of oblique contexts.

All of the principles, except for the collapse of the sense and denotation for canonical sense-names, are Fregean principles.[32] All are attractive, and can be independently motivated. I believe that they are very plausible components in any theory of the structure of thought. The principles of functional compositionality and functional decomposability of sense seem to me to be deeply embedded in computational psychological theories of thought. Senses are ways of thinking. Ways of thinking type-identify cognitive capacities. Cognitive psychology and philosophy need to attribute and account for cognitive capacities in ways that correspond to computational abilities, including inferential abilities. In order to do so, they need to specify particular representational abilities that enter into thinking in such a way that they determine the representational perspective and the computational capacities of the full propositional attitude. That is compositionality. Moreover, the attitude perspective and the computational capacities associated with the whole attitude need to be specified in such a way that one can recover specifications of the representational capacities that figure in it. That is decomposability. I think it plausible that any scientific language fit to specify such propositional attitudes for explanatory purposes in

[32] Frege supports the principles governing sense composition and decomposability in the following passages: Letter to Jourdain, *Philosophical and Mathematical Correspondence*, 79; *Wissenschaftlicher Briefwechsel*, 127; "Compound Thoughts" in *Collected Papers on Mathematics, Logic, and Philosophy*, ed. B. McGuinness (Oxford: Basil Blackwell, 1984), 390; *Kleine Schriften*, 378; O 36. Cf. also letters to Russell in *Philosophical and Mathematical Correspondence*, 149, 157, 158, 163, 165; *Wissenschaftlicher Briefwechsel*, 231, 239, 240, 245, 247; "On Concept and Object", in *Collected Papers*, 193; *Kleine Schriften*, 178; O 205; "On Sense and Denotation", in *Collected Papers*, 163, 166; *Kleine Schriften*, 148–149, 151; O 33, 37. Frege makes clear in his letter to Russell 12/28/1902 that he regards the natural-language shifting of sense and denotation of expressions in attributions of attitudes as an undesirable ambiguity. For a formal language he prefers the introduction of new signs, "though the connection with the corresponding signs in direct speech should be easy to recognize". This view indicates that he would have regarded a Method II language as an ideal language for thought. It also constitutes commitment to the substitution principle. The principle that free-standing sentences that have the same sense (express the same complete thought) have the same truth-value is implied by Frege's numerous discussions of thoughts as the bearers of truth or falsity. Cf. e.g. "The Thought", in *Collected Papers*, 353–354; *Kleine Schriften*, 344–345; O 60–61.

psychology ought to accept the compositionality and decomposability principles.

Substitutivity of identity is necessary in a metalanguage in which a truth-conditional semantics is to be carried out. Moreover, in the argument, substitution is applied only in contexts that do not correspond to oblique contexts.

The idea that free-standing sentences with the same sense have the same truth-value is, I think, also fundamental. Again it is crucial to remember that senses are ways of thinking—representational thought contents. I will discuss this principle at greater length than the others.

First, let us clear away a possible misunderstanding. We are not discussing the senses of free-standing (unapplied) sentences containing indexicals. Such sentences do not even fall under the principle, since the senses—that is, the ways of thinking—expressed by such sentences, apart from application in a context, do not have truth-values. Only when such sentences are supplemented by applications of the context-dependent devices in a context does one have a complete free-standing sentence that expresses a sense (way of thinking, a complete thought) that determines a truth-value.

The principle that free-standing sentences with the same sense have the same truth-value is, I think, an acknowledgment of a fundamental commitment, that is intuitive as well as theoretical. This is the principle that complete *thoughts* expressed by sentences, or by sentence-occurrences, have definite truth-values. (I lay aside non-denoting singular thought components and issues having to do with vagueness.) Complete propositional thoughts *are* truth conditions. They are true or false (or truth-valueless)—but not each, relative to some further parameter. In other words, propositional ways of thinking expressed by "complete" sentences are truth-bearers. I believe that they are the fundamental truth-bearers. Certainly, one of the fundamental intuitive facts about cognitive states is that some of them can be true or false, not merely true relative to one linguistic context, or mode of linguistic expression, and false relative to another.[33]

Suppose that one attempts to block the argument by denying the principle. One thereby disallows the transition from (e′) to (f′). Then one simply accepts that the sense of "<Opus 132> = <Opus 132>" is the same as the sense of "<Opus 132> = Opus 132". One might reason as follows: A sense may either determine itself, or it may determine its standard reference. Truth-value

[33] This claim does not entail Frege's view that ways of thinking are eternal or context-free. I think that some elements of senses (ways of thinking that happen to be expressed by language) are ineliminably individuated in terms of token representations—what I call "applications". Such representations are individuated in terms of occurrences in time, and are essentially dependent for their representational features and representational success (or failure) on context. On this point, see my "Reference and Proper Names", *The Journal of Philosophy*, 70 (1973), 425–439; "Demonstrative Constructions, Reference, and Truth", *The Journal of Philosophy*, 71 (1974), 205–223; "Belief *De Re*", *The Journal of Philosophy*, *op. cit.*; "Russell's Problem and Intentional Identity", in *Agent, Language, and the Structure of the World*, James E. Tomberlin, ed. (Indianapolis: Hackett Publishing Company, Inc., 1983), pp. 79–110; and "Five Theses on *De Re* States and Attitudes", *op. cit.*

can vary according to which referent the sense has. The referent determined by a sense is relative to a linguistic context.

I see no scientific or philosophical motivation for such a view. Even apart from regarding senses, with Frege, as ways of thinking, the idea that two complete (non-indexical), free-standing sentences with the same sense could have different truth-values is thoroughly unintuitive. It is also, as noted, theoretically unattractive inasmuch as it gives up any straightforward connection between senses and truth conditions.

But when one regards senses, as one should, as ways of thinking, the line of thought is more deeply unacceptable. I have already remarked on the absurdity of step (a′), which is a near-consequence of the view we are discussing. The view also entails that the way of thinking expressed by a singular term can be the same as the way of thinking expressed by a sentence. (The sense of "<Masterpiece(Opus 132)>" is supposed to be the same as the sense of "Masterpiece(Opus 132)".)[34]

The main considerations against this line are philosophical and scientific. Senses are ways of thinking. Where they are expressed by complete free-standing sentences, the ways of thinking are complete, propositional thoughts, or thought contents. Thought contents help type-identify propositional attitudes—psychological states. In an account of propositional attitudes—whether empirical-scientific, logical, or philosophical—canonical specifications of thoughts and thought components that are used in *attributions* of propositional attitudes are the fundamental way in which such attitudes and abilities are specified. Psychological laws or law-like generalizations can be expected to work off of such specifications. For purposes of logic, specification of thought contents and sub-propositional ways of thinking must bear a simple relation to propositional attitude attributions, if logic is to be used as a norm for thinking. So for purposes of understanding thought, as opposed to understanding more practically oriented natural-language ways of communicating about thought, we want a form of propositional-attitude attribution that centers on the nature of the attitudes themselves, and the nature of the propositional representational content that type-identifies the attitudes.

So discourse canonically specifying, and attributing, thoughts and ways of thinking is the most fundamental discourse about the natures of thoughts and ways of thinking. Suppose that one holds that in the most fundamental discourse about the natures of thoughts, the truth-value of a propositional representational thought content can be specified only relative to a linguistic context. Suppose one holds that in the most fundamental discourse about the

[34] In contrast to Frege, I think that the sense of a sentence (the way of thinking expressed by a sentence) and that of a singular term can never be the same. Frege thought that sentences and sentence-nominalizations could have the same sense. However, I believe that Frege's insistence on distinguishing customary sense from indirect sense suggests that he would have rejected the view that the sense of a sentence and the sense of a name of its customary sense could be the same.

natures of non-propositional ways of thinking that the *representatum* of a way of thinking can be specified only relative to a specification of an expression that expresses the way of thinking. Then one holds that the most basic features—the representational features—of thoughts and ways of thinking can be specified only relative to a linguistic context. Since thoughts and ways of thinking type-identify propositional attitudes, this implies that the most basic features of propositional attitudes can be specified only relative to a linguistic context. Any thought can be attributed at the base of a sequence of iterations. So to block the hierarchy in this way, one must hold this view for all thoughts. Such a view is committed either to the claim that all thought is language-relative (dependent for its most basic features, I would say its very nature, on relation to language) or to the claim that we have a peculiar inability to think about thought as it is. (Cf. note 8.)

There are, of course, philosophers who do maintain that all thought is dependent for its nature on language. This view seems to me to be incompatible with much that is known about both animal psychology and the development of language in human children. Arguments for the view seem to me to have been quite unimpressive.

This philosophical issue seems to me to be not quite as basic, for present purposes, as a scientific one. The view that in attributions of attitudes, one can specify the truth-value of a thought, or the *representatum* of a non-propositional way of thinking, only as being relative to a linguistic context, would make sciences of thought very *peculiarly* language-relative. Even if ways of thinking were all somehow metaphysically language-relative, I see no ground for holding that what they are about and what truth-values they have must be specified, in an empirical scientific theory or a logic, only relative to a particular linguistic expression within the language of the science. This would be an exceptionally strong species of language-relativity that would seem to hold in no other scientific domain. The point of an empirical science or a logic is to specify its subject matter in as general and context-free way as possible. Rejecting the principle that ways of thinking determine a definite *representatum* and complete thoughts have definite truth-values—in favor of the supposed language-relativity—would block normal scientific specification of the subject matters of psychology and logic.

I take the four principles to be fundamental to a logic and psychology of propositional attitudes. Unless one can give grounds for giving up one of the principles as part of these enterprises, one must take the hierarchy to be a fact embedded in the nature of thought. Thus blocking one or more of the principles within a particular language, or for the purposes of accounting for the structure of some possible or actual natural-language locutions, does not suffice to show that the hierarchy is not necessary to a correct understanding of thought. That is how Frege took it. I believe that he was almost surely right.

Let me, however, make a last remark about natural language. One could concede that a sense hierarchy does appear in the metalanguage, but hold that it is not used in accounting for oblique contexts in the object-language, allegedly the natural language. This move seems to me very questionable. It is quite unobvious why one should not be able to translate the hierarchy of sense-names present in the metalanguage back into the object-language, or rather into a natural language. Again, bear in mind that this hierarchy of sense-names is simply a way of specifying ways of thinking or cognitive perspectives. How could one deny that we are capable of thinking in these ways through the medium of natural language? I know of no bar to compositional translation of *these* (non-semantical) expressions between object-language (or natural language) and meta-language.

One can deny principles that lead to a hierarchy. But not only are all the relevant assumptions Fregean. They are well motivated and very plausible. They are especially plausible, it seems to me, as applied to a theory of thought.

Frege's approach to the language of ascription of propositional attitudes, and to thought about propositional attitudes, leaves out a lot. Frege ignores the fact that in many ascriptions we do not care about the exact way that an individual thinks about a subject matter. Sometimes we simply specify the subject matter. In ascriptions we often specify the subject matter of a person's thought in ways that we do not believe correspond at all to the way the person thought about the subject matter. In some of these cases, we "quantify into" that-clause constructions.

Frege's passing over this important and widespread phenomenon derives from his interest in an ideal scientific language for attributing propositional attitudes. An ideal language for this purpose would center on the nature of the attributed attitudes, fully specifying that nature—specifying the way that the relevant person actually thought. He was not interested in natural language ascriptions *per se*. He was especially not interested in attitude ascriptions insofar as the ascriptions are governed by pragmatic or contextual communicative concerns. He was interested in an ideal language for attributing attitudes *in order to specify what the attitudes are*. This involves specifying the representational perspective of the individual with the attitudes. I believe that this is not merely an old-fashioned interest. It remains an interest of scientific psychology and of any common-sense attribution that centers on conveying the nature of the attitude as fully and accurately as possible.

Frege's approach also does not address attribution of *de re* attitudes. Here I believe that his account of thought and sense is inadequate to explain the nature of the attitudes themselves. There are elements of *de re* propositional attitudes that are not timeless thought contents.[35]

35 Cf. my "Belief *De Re*" and "Five Theses on *De Re* States and Attitudes", and the Introduction, Part II.

Nevertheless, I believe that Frege's structural approach is, as far as it goes and given its aims, fundamentally on the right track. When one focuses not on linguistic meaning but on thought expressed and attributed by language—and one thinks of thought content as perspective on or way of thinking about a subject matter—the insight and power in Frege's approach tend to emerge. In the context of those aims, I believe that the postulation of a hierarchy of cognitive modes of presentation, or of ways of thinking, or of perspectives that are associated with propositional-attitude attributions, is tenable and probably superior to theoretical alternatives.

The hierarchy is motivated and entailed by basic principles that plausibly apply to thought content, propositional perspectives, ways of thinking. The relevant principles are those of the functional compositionality and decomposability of senses—of representational thought contents—the functional dependence of truth-value on the denotations of constituent linguistic parts (or the *determinata* of constituent representational contents), and the functional (determination) relation between senses and denotations. I believe that in view of the way that senses are formed and denoted through canonical names, the hierarchy is far from the bugbear that most philosophers have presented it as being. In itself, it is simply another example of the functional productivity of language and thought.

The sense hierarchy is, to my mind, particularly interesting in that it provides a new example of how subject matter (here, senses or thought-component concepts) can help individuatively determine the ways we think about a subject matter—even as the way we think about the subject matter also representationally determines the subject matter. The canonical senses or concepts take a cue from Russell: Understanding them requires understanding their denotations. But the sense hierarchy remains Fregean: A cognitive perspective is partial and is to be distinguished from what it is a perspective on.

Appendix I: Direct Pressure on Method I

In "Frege and the Hierarchy" and in the text of this Postscript, I discuss Method I mainly by considering its reliance on Method II for a semantical theory. I show how a hierarchy is induced for a Method II metalanguage for Method I. Here I discuss a more direct argument that suggests that on plausible principles, Method I itself is directly committed to a hierarchy.

It will be recalled that Method I allows systematic ambiguity of terms as between non-oblique contexts and oblique contexts. The point at issue is how far ambiguity extends. One might believe that there are only one or two levels of sense that are associated with embeddings of oblique contexts. In this

Appendix I want to give and discuss an argument for the view that Method I should be directly committed to an infinite hierarchy of senses to account for embedded contexts. The argument will be very similar to the last of the arguments that I give in the text, with one important alteration.

A term like "Opus 132" will be treated by Method I as having a customary sense when it occurs in a non-oblique context—a context like "Opus 132 = Opus 132". It will be treated as denoting this customary sense in an oblique context. The question at issue is whether there is a further sense, beyond the customary sense, expressed in ordinary unembedded oblique contexts. A one-level theory of sense gives a negative answer to this question. The argument that I give will be directed against this answer. Substantially the same argument can be given against any theory that stops the level of sense (for a given level of embedding) at a finite level. Thus an adaptation of the argument can be directed against a Method I theory that limits itself to only two levels of sense: the customary sense of an expression, and a further sense expressed by the expression in an oblique context. Such a theory would deny that a further (third) level is needed to deal with what sense is expressed in an embedding of an oblique context within a further oblique context. Similarly, an adaptation of the argument can be directed against a Method I theory that holds that there are only three levels of sense.

To make the argument easy to read, let us introduce a notation for the phrase "the sense of expression *E* in Context *C*". Let this expression be abbreviated by the two-place functional expression "S". Let "N" denote the class of non-oblique contexts. For example, "N" would denote a class that contains both singular term positions in the sentence "Opus 132 = Opus 132" standing alone (where "=" is read as "is identical with"). Let "O_1" denote the class of oblique contexts that are not embedded in a further oblique context. For example, "O_1" would denote a class that contains the position of "Opus 132" (occurring obliquely) in the sentence "Bela believes that Opus 132 is a masterpiece". I use angle brackets ("$< \ldots >$") as I do in the text, except that I include a marker for the context in which the term within the brackets is to be taken as occurring. Thus "<Opus132, O_1>" is a canonical name for the sense of the expression "Opus 132", as "Opus 132" occurs in an unembedded oblique context. "<Opus132, N>" is a canonical name for the sense of the expression "Opus 132" as "Opus 132" occurs in a non-oblique context.

Here is the argument:

(1) $S(\text{"Opus 132"}, O_1) = S(\text{"Opus 132"}, N) = <\text{Opus 132}, O_1> =$
$<\text{Opus 132}, N>$.

(1) simply records the one-level-of-sense view of the semantics of a Method I language. We will not use the latter two identities in (1) in the argument. I record them simply for clarity.

(2) S("Opus 132 = Opus 132", N) =
S("Opus 132", N) ^ S("=", N) ^ S("Opus 132", N).

(2) says, in effect, that the sense of a sentence occurring in a non-oblique context and asserting the self-identity of Opus 132 is decomposable into the senses that the component parts of the sentence express in a non-oblique context.

(3) S("Opus 132", N) ^ S(" = ", N) ^ S("Opus 132", N) =
S("Opus 132", N) ^ S(" = ", N) ^ S("Opus 132", O_1).

(3) follows from (2) and (1) by substitutivity of identity. Note that the substitution does not occur in an oblique context, or in a context that formalizes an oblique context.

Let "a" be a name of the customary sense of "Opus 132", and let "a" be capable of occurring in non-oblique contexts. Let "a" express (in non-oblique contexts) any sense one likes that is compatible with these explanations of its denotation or reference and of its grammatical behavior. Then

(4) a = S("Opus 132", N) = S("Opus 132", O_1).

(4) follows from the explanation of "a" and (1).

(5) S("Opus 132", N) ^ S(" = ", N) ^ S("Opus 132", O_1) =
S("Opus 132", N) ^ S(" = ", N) ^ a.

(5) follows from (4) and (3) by the substitutivity of identity.

(6) S("Opus 132", N) ^ S(" = ", N) ^ a =
S("Opus 132 = a", N)

(6) follows from (5) by the compositionality of sense. The sense (or thought) composed appropriately of the senses of the semantically relevant parts of a sentence is identical with the sense (or thought) expressed by the whole sentence.

(7) S("Opus 132 = Opus 132", N) = S("Opus 132 = a", N).

(7) follows from (2)–(6) by the transitivity of identity.

(8) Opus 132 = a.

(8) follows from (7) by the truth of self-identities and the principle that free-standing sentences with the same sense are materially equivalent.

(8) is absurd. It says in effect that Opus 132 is identical with the customary sense of "Opus 132". Since the argument can apply to any expression, and since it can apply to any account that stops at a finite level of senses expressed in successively embedded oblique contexts, it shows that in a Method I language, sense and denotation (or reference) collapse if the principles relied upon in the argument are accepted.

The principal premises used in the argument are again the decomposability of sense, the compositionality of sense, and the principle that free-standing sentences with the same sense have the same truth-value. There are two further background assumptions used in the argument.

One is that it is permissible in the Method I language to use functional terms (containing "S") that specify the senses of expressions relative to contexts, where the specifications are in accord with a given theory's account of what those senses are. Thus one is producing functional specifications of senses with complex singular terms that can occur in non-oblique contexts. I cannot see any ground to reject this assumption.

The other assumption is that it is permissible in the Method I language to introduce a name like "a" that can denote a sense but occur in a non-oblique context. I believe that this assumption is really dispensable in favor of the first assumption. It is just that the argument would require more apparatus to deal with embedding of quotation marks within quotation marks, or some analogous system for denoting expressions that denote expressions. I do not see that there is any principled reason for rejecting either of the assumptions.

I have claimed that the argument just given can be adapted to show that the attempt to stop the hierarchy at a finite level, in accounts of embedded oblique contexts, will on the relevant principles collapse the distinction between sense and denotation. Since these matters are complex, I will illustrate how the argument works at one higher level—against a view of Method I languages according to which customary sense and indirect sense are distinguished, but indirect sense is identical with doubly indirect sense. That is, the following argument reduces to absurdity the view that in a Method I language, there are only two levels of sense. It reduces to absurdity the view that the sense of an expression occurring obliquely in an embedded oblique context is the same as its sense in an unembedded oblique context (though distinct from its customary sense, which is denoted in an unembedded oblique context).

I will use iterated quotation marks that embed quotation marks, but I think it clear that these can be dispensed with, without affecting the argument. Here is the argument:

(1′) S("Opus 132", O_2) = S("Opus 132", O_1).

(1a′) S("Opus 132", N) ≠ S("Opus 132", O_1).

(2′) S("S("Opus 132", O_1) = S("Opus 132", O_1)", N) =
S("Opus 132", O_1) ^ S(" = ", N) ^ S("Opus 132", O_1).

(2′) follows from the decomposability of the sense of a sentence asserting that the indirect sense of "Opus 132" is self-identical.

(3′) S("Opus 132", O_1) ^ S(" = ", N) ^ S("Opus 132", O_1) =
S("Opus 132", O_1) ^ S(" = ", N) ^ S("Opus 132", O_2).

(3′) follows from (2′) and (1′) by substitutivity of identity.

Let "b" be a name of the sense of "Opus 132", as it occurs in non-embedded oblique contexts. That is, "b" names the indirect sense of "Opus 132". It names S("Opus 132", O_1). Let "b" be capable of occurring in non-oblique contexts. Let "b" express (in non-oblique contexts) any sense one likes that is compatible with these explanations of its denotation or reference and of its grammatical behavior. Then

(4′) b = S("Opus 132", O_1) = S("Opus 132", O_2).

(4′) follows from the explanation of "b" together with (1′).

(5′) S("Opus 132", O_1) ^ S(" = ", N) ^ S("Opus 132", O_2) =
S("Opus 132", O_1) ^ S(" = ", N) ^ b.

(5′) follows from (3′) and (4′) by substitutivity of identity.

In an unembedded oblique context "Opus 132" denotes only its customary sense and expresses its indirect sense (S("Opus 132", O_1)). Let us introduce a term "B" that, in a non-oblique context, expresses the sense that "Opus 132" expresses in an unembedded oblique context and denotes the sense (the customary sense of "Opus 132") that "Opus 132" denotes in that same unembedded oblique context. So "B" denotes S("Opus 132", N). That is, "B" has the same semantical characteristics in a non-oblique context that "Opus 132" does in an unembedded oblique context. Then

(6′) S("Opus 132", O_1) ^ S(" = ", N) ^ b =
S("B = b", N).

(6′) follows from (5′) by the compositionality of sense and the explanation of the term "B".

(7′) S("S("Opus 132", O_1) = S("Opus 132", O_1)", N) = S("B = b", N).

(7′) follows from (2′)–(6′) by the transitivity of identity.

(8′) B = b.

(8′) follows from (7′) by the principle that sentences (standing alone) that have the same sense also have the same truth-value.

(8′) contradicts (1a′). Thus again, on plausible principles, a Method I language will collapse the hierarchy altogether if it identifies an expression's sense with its denotation at any level of embedding of oblique contexts.

The plausible principles are the decomposability of sense, the compositionality of sense, and the principle that sentences (standing alone) with the same sense have the same truth-value.

As in the first argument in this Appendix, there is one additional assumption: We can introduce into Method I languages names or descriptions that in non-oblique contexts express the senses and denote the denotations that other expressions express and denote in oblique contexts. I see nothing in the conception of Method I languages, or in the conception of Method I languages that I am criticizing, that prevents one from introducing such expressions.

What these arguments show is that the substitutivity of coextensive expressions that differentiates Method II from Method I languages, and which makes Method II languages superior as a basis for giving a truth theory, is not fundamental to yielding the arguments for the hierarchy. More basic is having expressions that can appear in non-oblique positions but which mimic the semantical behavior of expressions that appear in that-clauses. They must denote and (after the first-level) express the senses of expressions that appear in that-clauses. In a Method I language in which a term can occur either non-obliquely or obliquely at various levels of embedding, the argument simply requires that there be other expressions that have the same semantical characteristics—have the same sense and denotation—as occurrences of terms in oblique position. A scientific language that has any chance of describing the structure of sense or thought content as expressed in natural language must have such terms.

Thus on plausible principles, Method I languages generate a hierarchy just as surely as Method II languages do.[36]

Appendix II: Existential Generalization on Embedded Positions

In the main text of the Postscript, I discussed existential generalization on

(2) Igor believes that Bela believes that Opus 132 is a masterpiece.

(2) entails

(2Exp) There is something, namely that Opus 132 is a masterpiece, that Igor believes that Bela believes,

which entails

(2EG) There is something that Igor believes that Bela believes.

[36] In unpublished work Nathan Salmon has derived the hierarchy from different principles applicable within a Method I language. For example, he uses the shift principle that the indirect reference of an expression is the customary sense of the expression together with principles governing the identity between the reference of an embedded expression and the indirect reference of an expression to which a shift operator expression (such as the "that" in a that-clause, or "the thought that") is applied but at one lower level of embedding. Thus the key principles are those governing shifts within a Method I language. I believe that Frege clearly accepted the principles that Salmon relies upon, and that these further proofs of the hierarchy are interesting both in capturing Frege's thinking and in themselves. But I believe (and Salmon agrees) that the principles of functional composition and decomposition that I have centered on are more basic in Frege's thinking. And I believe that Frege obviously preferred a Method II language as an ideal language for attributing propositional attitudes. (Cf. again the letter to Russell 12/28/1902.)

I formalized (2) as

(2a) Believes (I, <Bela>^<Believes>^C(<Opus 132 is a masterpiece>)).

I formalized (2Exp) as

(EXP) (∃y)(y = <Opus 132 is a masterpiece> & Believes(Igor, <Bela>^<Believes>^ C(y))).

(2a) also formally entails

(EG) (∃y)(Believes(Igor, <Bela>^<Believes>^ C(y))).

I believe that (EG) formalizes (2EG).

It is worth nothing that (2a) also formally entails

(EG′) (∃z)(Believes(Igor, <Bela>^<Believes>^ z)).

Here it may be tempting to raise an objection to the account that I have given. One might be tempted to hold that in view of the fact that (EG) and (EG′) both follow from (2a), it is a consequence of (2a)'s formalizing (2) that I am committed to (2)'s entailing that Igor believes Bela believes at least two things.

I believe that this objection is mistaken. This Appendix will explain why. The larger point of the explanation is to elicit a better understanding of the hierarchical account, and to explore its relation to English. I believe that (EG) and not (EG′) formalizes (2EG).

As noted, (EG) and (EG′) are both entailed by (2a). It is worth noting that there are yet *further existential generalizations in the offing*. In the first place, there will be more existential generalizations (EG″), (EG‴) . . . with further embeddings ("Arnold believes that Igor believes that Bela believes that Opus 132 is a masterpiece"), and further iterations of "C" in the formalizations. In the second place, there are positions in the structured root proposition-name that denote senses that are components of the denoted proposition or thought content (e.g. the position of "<Opus 132>" in the structured "<Opus 132 Masterpiece>", which really has the form "<Opus 132> ^ <Masterpiece>"). From these positions one will have even more quantifications that follow from the initial formalization (2a).

I think that the first of the quantifications, (EG), is what formalizes the English quantification

(2EG) There is something that Igor believes that Bela believes.

Only (EG) captures the inference to (2EG) from

(2) Igor believes that Bela believes that Opus 132 is a masterpiece.

I believe that I can say why (EG) and not (EG′) formalizes (2EG). I will begin by explaining why (EG) and not (EG′) tracks the inference from (2) to (2EG). I will come back later to discussing the interpretation of (EG′).

The reason begins with this observation: (2EG) follows from (2) inasmuch as Igor is characterized in (2) as believing that Bela believes a certain *proposition* or *thought content* (namely, that Opus 132 is a masterpiece). That is, we got (2EG) by first reasoning from (2) to

(2Exp) There is something, namely that Opus 132 is a masterpiece, that Igor believes that Bela believes,

and then dropping the conjunct that consists in "namely that Opus 132 is a masterpiece".

Perhaps it is possible for Igor to be so odd that he believes that Bela believes things other than propositions. Perhaps—although this is at best questionable—Igor could believe that Bela believes a fire engine, or a stone. Perhaps (again questionably) Igor could believe that Bela believes the sense of a singular term, a sense which is not a thought content or proposition. Bela cannot *in fact* believe a fire engine or the sense of a singular term. What might be disputed is whether Igor could believe he can. It would be easier on us all if we could just declare that no one can believe that anyone believes something other than a proposition or a thought content. Perhaps then there would be fewer entailments to worry about! But let us proceed on the assumption that it is possible for someone erroneously to believe that someone believes something other than a proposition or thought content.

However, the only thing that it is guaranteed that Igor believes that Bela believes, *given the truth of the sentence (2)* , is a proposition or thought content—namely, the thought content that Opus 132 is a masterpiece. There are *not* two things that Igor believes Bela believes in virtue of the truth of (2).

I think that the same is true of (2a). (2a) does *not* entail anything that commits us to saying that Igor believes that Bela has another belief besides the belief that Opus 132 is a masterpiece. The only relevant term that denotes a proposition or thought content that Igor could believe Bela believes in my formalization (2a) is "<Opus 132 is a masterpiece>" – *not* "C(<Opus 132 is a masterpiece>)". The latter does not denote a proposition or thought content. It denotes the sense of the canonical name "< Opus 132 is a masterpiece >". (This latter canonical name denotes the customary sense of "Opus 132 is a masterpiece".) So, even if we grant that it is possible, it would be *very* odd of Igor to believe that a non-proposition is believed by Bela. And it is certainly not guaranteed or implied by (2) or (2a) that he believes that Bela believes any such thing. So if we export "C(<Opus 132 is a masterpiece>)", we are not exporting a name of some proposition that Igor believes Bela believes. It is wrong to think that the canonical name (or designator) names anything that Igor believes that Bela believes. I will come back to how we should read such exportations.

Both in English grammar and in formalizations of English, one should keep track of the level of entity (and within sense levels, the type of entity—e.g. sense of a singular term, sense of a sentence, and so on) that is being quantified over. One has to do some of this anyway to distinguish quantifying into that-clauses (in effect, all the way down to the bottom level of reference or denotation) from quantifying onto them (quantifying over a sense or mode of presentation). Thus, in a full theory of quantification, one should mark the variable (here "y") that "C" applies to as a variable that goes with the name of a propositional thought content—as distinguished from the name of a customary sense of a name, and as distinguished from the name of a customary sense of a predicate.[37]

We have the following situation: (2EG) follows from (2). (2a) formalizes (2). Both (EG) and (EG′) follow from (2a). So both (EG) and (EG′) should follow from (2). But (EG′) does not generalize on a second belief of Igor's, a belief over and above the belief that Opus 132 is a masterpiece (contrary to the tempting objection that we began with). So generalization on the position of "C(<Opus 132 is a masterpiece>)" in (2a) is not what is going on in the English (2EG). (2EG) is *not* true in virtue of a generalization on "C(<Opus 132 is a masterpiece>)". For this would take the relevant singular basis for the existential generalization to be the sense of a singular term, not a thought content—the sense of a sentence.

I stipulate that Igor does not believe that Bela believes the sense of a singular term. Igor attributes only propositional beliefs to Bela.

Another way of putting all this is that we cannot get to a formalization of (2EG) by exporting "C(<Opus 132 is a masterpiece>)" from (2a) to get

(EXP′) (∃z)(z = C(<Opus 132 is a masterpiece>) & Believes(Igor, Bela> ^ < Believes > ^ z)),

taking this to be a formalization of

(Weirdo) There is something, namely the canonical sense of the name "that Opus 132 is a masterpiece", that Igor believes that Bela believes

and then dropping the first conjunct of the existential generalization to get

(EG′) (∃z)(Believes(Igor, <Bela> ^ <Believes> ^ z)),

taking this to formalize

(2EG) There is something that Igor believes that Bela believes.

There is no analogy here to first exporting "that Opus 132 is a masterpiece" from (2), then formalizing this exportation by (EXP), and then simplifying (EXP) to (EG).

In view of the fact that the "z" in (EG′) does not trace back to the name of a proposition, it is not made true by anything Igor believes that Bela believes. Moreover, (Weirdo) is *not a correct reading* of (EG′). For "the canonical

[37] Of course, Church's Logic of Sense and Denotation has this feature.

sense of the name 'that Opus 132 is a masterpiece' " is a definite description and "C(<Opus 132 is a masterpiece>)" is a canonical name or canonical designator—not a definite description. Our canonical names represent specifically oblique occurrences of expressions in natural language, most prominently such occurrences in ordinary or in embedded that-clauses. "C" specifically marks levels of embedding in canonical names of modes of presentation or representational content.

Just to fix this last point, let us consider formalizations of

(2Weird) Igor believes that Bela believes the fire engine

(2Weirdo) Igor believes that Bela believes the canonical sense of the name "that Opus 132 is a masterpiece".

I believe that one can take each of these as having a quantifying-in reading and an oblique occurrence reading, if they have any readings at all. Let us focus on the oblique readings. The propositions that Igor believes are *expressed* by the sentences

(2Weird -) Bela believes the fire engine

(2Weirdo -) Bela believes the canonical sense of "that Opus 132 is a masterpiece",

where the direct objects occur obliquely. Suppose that the terms in direct-object position denote their customary senses. Suppose that we allow first-level canonical names to formalize those occurrences and to denote the customary senses:

<the fire engine>

<the canonical sense of "that Opus 132 is a masterpiece">.

Then in formalizations of (2Weird) and (2Weirdo) we must introduce canonical names that denote the senses of these canonical names, as those senses are expressed in (2Weird -) and (2Weirdo -). Thus we formalize (2Weird) and (2Weirdo) as

(2Weird - f) Believes(Igor, <Bela> ^ <Believes> ^ C(<the fire engine>))

(2Weirdo - f) Believes(Igor, <Bela> ^ <Believes> ^ C(<the canonical sense of "that Opus 132 is a masterpiece">)).

I believe that these formalizations convincingly suggest two things. One is that if one keeps track of whether a canonical name formed with "C" derives from a root canonical name of a complete propositional thought content, or a root canonical name of a sense other than such a content, one will never get formalizations in embedded cases that confuse the two, or that are inspecific as between the two.

The other thing suggested by the formalizations is that canonical names enter initially as formalizations of occurrences of *other* expressions of the natural language that occur obliquely. On Method II, which we are following, *no canonical name can itself occur obliquely*. The first-level canonical names (which I have been formalizing with angle brackets) will formalize all root weird occurrences. (Cf. the formalizations of (2Weird -) and (2Weirdo -).) "C" comes into play *only* in formalizations of embeddings—usually embedded that-clauses, but perhaps also embedded attributions of Weird beliefs. I believe that *in no case* is it true that such formalizations allow a canonical name which includes "C" to formalize a root direct object of an embedded occurrence of "believe"—whether the object denotes a proposition or not. Thus (EG′), understood as quantifying onto a place that had contained a canonical name that includes "C" does not formalize any sentence like (2Weird) or (2Weirdo). I conjecture that wide-scope quantifications that formalize natural-language quantifications on the direct object of the final or root direct object always leave the "C" (possibly a string of iterated "C"'s) in place (i.e. not exported), as (EG) does. I will soon discuss how to read back into English formalizations that do export canonical names that include "C".

There are surely further things to be said of

(2EG) There is something that Igor believes that Bela believes.

(2EG) should follow even from Igor's believing that Bela believes a fire engine or believes the sense of a singular term. Our formalizations follow this course. For example, (EG), which does follow from (2Weird) and (2Weirdo), is a formalization of (2EG). As I said, I believe that the English grammar as well as the formal theory should leave a trace of the fact that here (2EG) derives from an attribution by Igor of a non-propositional belief to Bela.

The key point for present purposes is that (EG′) should not be the formalization of the generalization (2EG) that follows even from (2Weird) and (2Weirdo). It is (EG) that is the relevant formalization of the generalization (2EG) that follows from (2Weird) and (2 Weirdo).

So an issue has become: what *other* English sentence might be formalized by (EG′)? Let us first consider where the existential generalization in (EG′) came from. Let us think about how we would read the exportation of "C(<Opus 132 is a masterpiece>)" from (2a). The formalized exportation is

(EXP′) (∃z)(z = C(<Opus 132 is a masterpiece>) & Believes(Igor, <Bela> ^ <Believes> ^ z)).

This can be glossed in English:

(Engl-Exp′) There is a (canonical indirect-sense-level) way of thinking about the thought that Opus 132 is a masterpiece that Igor's belief utilizes in attributing that thought (that Opus 132 is a masterpiece) to Bela as a belief.

One could get from this reading down to this English gloss on (EG′):

(Engl-EG′) There is a way of thinking about a thought that Igor's belief utilizes in attributing a thought to Bela as a belief.

Or, more briefly,

(Engl′-EG′) There is something that Igor's belief utilizes in attributing a belief to Bela.

Could this be the English sentence that (EG′) formalizes?

One might protest that these glosses have "extra" words—"way of thinking", "thought", "utilizes", "attributing", that have no counterparts in the formalizing sentences (EG) or (EG′). The fact that there isn't a *smooth* reading in English of (EG′) is not, I think, a difficulty.

I believe that the situation with the English in this case is broadly analogous (but perhaps not ultimately as deeply interesting) as the problem about expressing second-order quantification in English. From the natural formalization of "Sally found Bill"—F(s, b)—, one can infer

$(\exists F)F(s, b)$.

One can put this in English as something like

Sally bore some relation to Bill.

This sounds more "English", as well as more genteel, than "Sally F'd Bill, for some F".

One might reply, "But then the verb is 'bore' ", or "But then the relation is <u>bore</u>". Well, yes. But it does not follow that we cannot do enough meta-talk about how to understand the English that it is clear that we are understanding words like "bore" as syncategorematic, or as indicators of the second-order quantification. Surely one can express the second-order quantification in English. So it seems that we should be able to explain an understanding of the English according to which we have a second-order quantification over a two-place relation (in "Bill bore some relation to Sally") rather than a first-order quantification over an argument or *relatum* in a three-place relation <u>bore</u>. I think that one *can* engage in such meta-talk and that English can express the second-order locution, though it does not do so smoothly on its surface.

In the case of the "excess" quantifications in the embedded that-clause cases (of which (Engl′-EG′) is an example), it is not a matter of second-order quantification. But I think getting the readings in English of the formalizations will force "extra" words into the English in an analogous way. Here the syncategorematic words are ultimately "utilizes" and "attributes".

I conjecture that any view that both treats that-clauses as *singular terms* with *structure* and treats that-clauses in natural language as having *oblique* positions (i.e. any view that has any chance of being correct as an account of English) will throw up quantifications that cannot be smoothly read back into English, but which are true and are credibly expressible in English.

Part II

Sense and Cognitive Value

5 *Sinning Against Frege (1979)*

Fregean *Sinn* has been provocative, seminal, and prolific. But ever since it was propagated in English speaking philosophy, it has been widely misunderstood. Recent condemnations of *Sinn*—from Searle and Wittgenstein to Kripke and Donnellan—have to a significant degree rested on misunderstanding. My mission here is primarily historical. It is to trace the misunderstanding, and right some of the historical wrongs. I will not try to redeem Frege from all transgression, nor will I count *Sinn* a virtue. But I believe that better acquaintance with *Sinn* is a precondition for successfully eschewing it.

The basic misunderstanding is the identification of Frege's notion of *Sinn* (sense) with the notion of linguistic meaning. The misunderstanding is an easy one to fall into for two reasons. For one, the term "meaning" has always been vague, multi-purposed, and to some extent adaptive to the viewpoint of different theories. Pressing the term into service to characterize Frege's notion has seemed harmless enough, as long as it is made clear that the notion is restricted to an aspect of meaning relevant to fixing the truth-value of sentences. A second reason for the misunderstanding has been that Frege did not lavish any considerable attention on the area in which the differences between sense and the ordinary notion of meaning are clearest—context-dependent reference.

Although the differences between meaning and sense are easiest to notice with indexicals (including proper names), the distinction issues from the fundamental cast of Frege's work, a cast discernible throughout his career independently of issues about indexicals. Baldly put, Frege was primarily interested in the eternal structure of thought, of cognitive contents, not in conventional linguistic meaning. He pursued this interest by investigating the structure of language, and much of his work may be seen as directly relevant to theories of linguistic meaning. But the epistemic orientation of his theorizing leads to a notion of sense with a different theoretical function from modern notions of meaning.

I am grateful to the John Simon Guggenheim Foundation for its support, and to the referee [of *The Philosophical Review*] for comments. Scholarly tradition credits the moralistic tone of Fregean research to Paul Benacerraf.

The sign "O" in citations marks the pagination in the original publications of the articles by Frege.

I

Why is it a mistake to identify Fregean sense with meaning? The grounds for avoiding the identification are most evident in "The Thought," where Frege discusses various indexical expressions. Frege argues that the thought expressed by an indexical sentence, which he identifies with its sense, may remain the same even as the sentence is changed.

(*A*) If a time indication should be made in present tense, one must know when the sentence was uttered to grasp the thought correctly. Thus the time of utterance is part of the expression of the thought. If someone wants to say today what he expressed yesterday using the word 'today', he will replace this word with 'yesterday'. Although the thought is the same, the verbal expression must be different, to compensate for the change of sense which would otherwise be brought about by the different time of utterance. The case is the same with words like 'here' and 'there'. In all such cases, the mere wording, as it can be written down, is not the complete expression of the thought—but one further needs for its correct apprehension the knowledge of certain circumstances accompanying the utterance, which are used as means of expressing the thought. Fingerpointings, gestures and glances may belong here too. The same utterance containing the word 'I' will express different thoughts in the mouths of different people, of which some may be true and others false.[1]

According to Frege, when the context of utterance shifts, the sense of sentences containing "I," "here," "there," "yesterday" or "today" can also shift. Since the senses of other words of the sentence need not shift, the shift in sense is associated with the indexicals. (See note 1; esp. *Nachgelassene Schriften*, 146, where this point is more explicit than in (*A*).) It is not important to our interpretation whether one sees the indexical as "expressing" a sense in the context (Frege does not use this locution), or whether one sees the sense that is subject to contextual shift as expressed (or indicated) by the "circumstances accompanying the utterance," the person's demonstrations, and so forth. I hedge this point by writing of the sense associated with

[1] Gottlob Frege, *Logische Untersuchungen* (Göttingen: Vandenhoeck und Ruprecht, 1966), 37–38. Translations of this article throughout are mine. See "The Thought," in Klemke (ed.) *Essays on Frege*, (Urbana, Ill.: University of Illinois Press, 1968), 516–17; O 64. The translation by A. M. and M. Quinton contains an ungrammatical English sentence in passage (*A*) and is in certain other minor respects not quite as literal as the present one. For example, in the third sentence, their translation (without justification) reads "must" instead of "will." Page references in the text are to the Klemke volume. Most of the views, and even some of the particular phrasings, of "The Thought" (1918–19) may be found in two unpublished introductions to what was apparently planned to be a textbook entitled "Logik." The first was written sometime between 1879 and 1891; and the second, of somewhat greater length, is dated 1897. See *Nachgelassene Schriften*, Hermes, Kambartel, and Kaulbach, eds. (Hamburg: Felix Meiner, 1969), 1–8; "Logik" (zwischen 1879 und 1891); 137–63; "Logik" (1897). In the latter fragment Frege makes it even clearer that the sense associated with indexicals shifts: "Words like 'here', 'now' achieve their full sense always only through the circumstances in which they are used . . . and the same sentence does not always express the same thought, because the words require supplementation to yield the complete sense, and . . . this supplementation can be different according to circumstances" (p. 146).

the indexical in the context. Clearly Frege holds such senses may shift with context.

Do the meanings of indexical expressions shift? The most natural answer to this question is clearly "no." The relevant expressions are each governed by a single linguistic rule and have a single context-free dictionary entry. In learning the meanings of these words, one comes to know how to use and understand the words regardless of what occasion arises. Given a context and the meaning of the expression, the referent can usually be determined. Thus on the most natural construal of the notion of meaning, sense and meaning must be distinguished.

Frege also clearly thinks that the thought or sense expressed in an indexical utterance *can* be the same as that expressed in another utterance containing an indexical with a different meaning. As is stated in (*A*), "yesterday" and "today," used in appropriately different contexts, can be employed to express the same sense. Here, sense remains constant while meaning shifts. So the linguistic meaning of indexicals need not even be *part* of the sense associated with them in a given context. The indexically identified referent may be presented in thought independently of the particular mode of indexical expression used to communicate the thought. This should not be surprising in view of the eternal context-free nature of senses, a feature we shall discuss later.

In sum, a single indexical expression ("today") may be associated with different senses on different occasions; and indexical expressions with different meanings ("yesterday" and "today") may, in their respective contexts, be associated with the same sense. (*May* be: as I shall next argue, the fact that they have the same referent in their respective contexts, does not guarantee that they are associated with the same sense; in most cases, one would expect them to be associated with different senses, since the thinker's epistemic perspective is likely to differ.) Sense seems to vary independently of the meaning of indexicals.

Some philosophers claim that in one sense of "meaning," the meaning of an indexical expression is its referent. Now it is clear that the referents of "I," "here," "today" and so forth shift with the context. Could we count indexical expressions with the same reference as always making the same contribution to the thought expressed on the respective occasions? To put it another way, could we identify sense with reference ("meaning") for these indexical expressions?[2]

[2] Michael Dummett, in *Frege: Philosophy of Language* (London: Duckworth, 1973), 384, misinterprets Frege in this way, basing the interpretation on passage (*A*). Actually Dummett only identifies thoughts expressed by $A(s_1 \ldots s_n)$ and $A(t_1 \ldots t_n)$, where $s_1 \ldots s_n$ and $t_1 \ldots t_n$ are indexical expressions with the same respective referents. He says nothing about the sense associated with the indexicals. But the difference between this interpretation and the one we discuss is irrelevant to the points we make. Dummett's interpretation is ably criticized by John Perry, "Frege on Demonstratives," *The Philosophical Review*, 86 (1977), 474–497, who cites the third and (in a different form) the second of the reasons against Dummett's interpretation given below. Perry clearly distinguishes

"No" again. There are two textual reasons and one systematic reason for scouting this reading. In the first place, the passage (*A*) does not clearly support the interpretation. Frege says that the time of utterance is part of the *expression* of the thought—he does not say that it is a component of the thought. Surely Frege would have announced and explicated the identification of the sense of an indexical (in a context) with its referent if he had believed in such an identification. Both the sense and referent of indexical expressions shift with context, but this is because the sense is the epistemic basis for determining the referent—not because it *is* the referent.

A second reason for rejecting the identification of sense and referent is that in the pages immediately following (*A*), Frege clearly indicates that "I" may be used with different senses (giving rise to different thoughts) even though it is applied to the same person ("The Thought," 519; O 66). Moreover, in the same section Frege treats proper names as having different senses while applying to the same person. If Frege had envisioned a sharp distinction between proper names and certain indexicals on this matter, he would have reported his vision.

In this section, Frege repeats his view that sense accounts for the way a referent can be presented indexically in distinct ways. This constitutes the systematic ground for not identifying sense and referent. It is possible to believe what is expressed by "Today is Friday" and (without in any ordinary sense changing one's mind) doubt what is expressed by "Yesterday was Friday," even though "yesterday" and "today" (in their different contexts) pick out the same day. Similarly, for most of the other indexical expressions Frege mentions. But Frege's primary motivation for introducing sense was to account for differences in cognitive value. Thus the thought expressed by the different utterances may be different, as will the senses associated with the indexicals. There is no reason to think that when he came to indexicals, Frege forgot his own ground for postulating senses.[3] Thus the sense of an indexical

Frege's notion from both meaning and reference, though he is less charitable in his renunciation of *Sinn* than I think appropriate. In particular, Perry writes as if Frege made a *mistake* in identifying the senses of sentences, which Perry thinks are naturally taken to be meanings (or what he calls roles), with thoughts. I think this view reflects the picture of Frege as a theorist primarily of meaning rather than of thought—a picture I shall argue is distorted. Given Frege's original purpose in introducing sense—to account for differences in possible belief—given his consistent explication of the notion in terms of a mode of presentation, and given his explicit disavowals of concern with language, it seems better to see Frege's notion of thought as explicating what he *meant* by "sense." Frege's mistake (for our present purpose) lies in his account of the nature of thoughts (context-free and "complete in every way"). These remarks are elaborated below. For a discussion of the distinction between meaning and sense and a criticism of Frege in a less historical and more constructive setting, see my "Belief *De Re*," *The Journal of Philosophy*, 74 (1977), sect. IV. I think there are deep problems with Frege's account other than those detailed in "Belief *De Re*" and in the present article. But discussing them would carry us beyond our present purpose.

[3] See the opening section of "On Sense and Reference", in *Translations of the Philosophical Writings of Gottlob Frege*, ed. Geach and Black (Oxford: Blackwell, 1966). Citations of this article in the text will be to this volume. In a letter of 1919 (at roughly the same time as the publication of "The Thought") to the historian Darmstaedter, Frege distinguishes sense and reference by the usual appeal to

should not be identified with its meaning, whether one construes "meaning" in a more or less ordinary way, or as amounting to reference.

The same sort of point can be made about Frege's view of the sense of proper names. In "On Sense and Reference" Frege writes

(*B*) In the case of an actual proper name such as 'Aristotle' opinions as to the sense may differ. It might for instance be taken to be the following: the pupil of Plato and teacher of Alexander the Great. Anybody who does this will attach another sense to the sentence 'Aristotle was born in Stagira' than will a man who takes as sense of the name: the teacher of Alexander the Great who was born in Stagira. So long as the reference remains the same, such variations of sense may be tolerated, although they are to be avoided in the theoretical structure of a demonstrative science and ought not to occur in a perfect language. ["On Sense and Reference," 58; O 28n]

Shortly thereafter Frege writes, "To every expression belonging to a complete totality of signs, there should certainly correspond a definite sense; but natural languages often do not satisfy this condition, and one must be content if the same word has the same sense in the same context." The implication is clearly that the senses associated with proper names and other indexical constructions shift with context. In "The Thought" (pp. 516–518), Frege makes similar remarks emphasizing the variability of the sense of a proper name for different users and in different contexts. In this respect Frege treats names and indexicals in the same way.[4] Thus for reasons similar to those

the paradox of identity and writes: "When an astronomer asserts something of the Moon, the Moon itself is not part of the expressed thought" (*Nachgelassene Schriften*, 275, see also *Translations*, 64; O 35).

[4] It is less clear what Frege thought the differences are between names and (ordinary) indexical expressions. He emphasizes that in a sense names are not part of a natural language ("The Thought," 517; O 65); one can be fully competent in the language but fail to associate with a name the sense associated by the speaker in the context. But the same could be said of many uses of indexicals. (See further discussion below.) Two papers which are not historically oriented, but which take up positions on the role of indexicals and names in expressing thought, as distinguished from effecting communication, that are very broadly similar to Frege's are Hector-Neri Castañeda, "On the Philosophical Foundations of the Theory of Communication: Reference," *Midwest Studies in Philosophy*, 2 (1977), esp. 172–173; and Brian Loar, "The Semantics of Singular Terms," *Philosophical Studies*, 30 (1976), 353–377. See also n. 22 below.

Relevant to the interpretation of Frege just proposed, Ruth Marcus writes in a recent review (*The Philosophical Review*, 87 (1978), 503):

There is perhaps another theory of sense to be culled from some Fregean texts. . . . On that (alternative?) reading the sense of a term is whatever is grasped or understood by a speaker on a particular occasion of use and may vary from occasion to occasion as well as from speaker to speaker. . . . To see this as *the* Fregean view runs counter to Frege's anti-psychologism and his belief in the "common stock of thoughts" in a community of speakers. What would become of the Fregean slogan "To give the meaning is to give the truth conditions"?

All that Frege writes about indexicals and proper names makes it clear that the view described is not an alternative reading, but is the only correct reading of his view of *these* terms. Frege defends his anti-psychologism largely through his treatment of thoughts, and of logico-mathematical objects and laws, as ontologically independent of minds. The anti-psychologism is fully compatible with his views about indexicals and proper names, which for him did not belong to a language of pure thought in any case. Nor is there an inconsistency between these views and there being among thinkers a common stock of thoughts. People may have common thoughts, whether indexically or non-indexically expressed. Strictly speaking, the cited Fregean slogan is not to be found in Frege. He showed by example that one should analyze component senses with an eye to their role in fixing truth or falsity. Such analysis given earlier, the sense of a proper name should not be identified with its meaning.

It is sometimes held that names do not have a meaning. Justifications for this view usually allude to the point that different competent speakers may not "understand" each other's use of a name. But Frege himself makes this point in *explaining* the notion of sense (see (*B*)). So the point cannot responsibly be used against him. Rather, Frege's notion of sense, as applied to proper names, is such that a competent speaker may not catch on to the sense of another competent speaker's use of a name. Insofar as this cannot be said of the meaning of a name, if any, sense and meaning must be distinguished. The cognitive value of a name (at a given occurrence) may well be idiosyncratic.

Sometimes it is held that the meaning of a proper name is its bearer. But it is clear, for reasons we adduced in connection with ordinary indexical constructions, that sense is not identifiable with meaning on this construal. Moreover, if there is a meaning associated with a name like "Aristotle" that is mastered merely by learning how such names are used in the language—when one learns what it is to be an Aristotle[5]—then sense and meaning in *this* sense are not identifiable: Sense shifts with context; meaning does not.

These passages dealing with the context dependence of indexical constructions, including proper names, are but the most obvious signs of the cognitive orientation of Frege's notion of sense. That orientation dominates his introduction of the notion in "On Sense and Reference." The paradox of identity, whose discussion opens the essay, is a problem about information expressed through language. That problem can be shown to resist attempts to solve it by reference to differences of meaning. Two indexical identity-sentence occurrences with the same component referents and even the same linguistic meaning may differ in informational or cognitive value, and may be used to express different beliefs. For example, "this is identical with this," or "Bertrand is the same person as Bertrand," may be informative in one occurrence and trivially true in another. (See "Belief *De Re*," note 2 above.) Moreover, Frege's solution to the problem in terms of different senses is expressly cognitive in character. As noted earlier, sense is explicated as containing the "way of being given" (*Art von Gegebensein*)—the mode by which an object is presented in thought ("On Sense and Reference," *Translations*, 57; O 26). This association is maintained throughout his writings. The senses of nonindexical sentences, or sentences used in a context, are thoughts. Thus senses are, or are components of, abstract thoughts (*Gedanken*, thought contents). (See "On Sense and Reference," 59, 62–63; O 29, 33; and *Nachgelassene Schriften*, 275; "Aufzeichnungen für Ludwig Darmstaedter"

applies to thoughts whether contextually expressed or not. It should be noted that Marcus is apparently arguing against taking the view to be Frege's view of the sense of *all* terms (although her discussion does key on proper names and does not distinguish meaning and sense). She is clearly right in holding that there is no basis for taking Frege to have applied the view in question to all terms.

[5] See my "Reference and Proper Names," *The Journal of Philosophy*, 70 (1973), 425–439.

(1919).) Thoughts may be expressed and apprehended through language. But they are ontologically and conceptually independent of language and of human agents. (See "The Thought," 533–534; O 75–76; and *Nachgelassene Schriften*, p. 146; "Logik" (1897).)

It is well known that Frege more than once disavowed primary concern with language. These disavowals have been taken to be a sign of weakness, a sort of fallback position utilized when difficulties threatened. There may be some truth in this point. But the basic import of Frege's remarks is best grasped by taking him at his word:

> If it is one of the tasks of philosophy to break the domination of the word over the human mind by laying bare the misconceptions that through the use of language almost unavoidably arise concerning the relations between concepts and by freeing thought from that which only the means of expression of ordinary language, constituted as they are, saddle it, then my ideography . . . can become a useful tool for the philosopher.[6]

> I have to content myself with presenting the reader with a thought, in itself immaterial, shrouded in sensible linguistic form. The metaphorical aspect of language thus presents difficulties. The sensible always breaks in and makes expression metaphorical and so improper. So a battle with language arises and I am compelled to occupy myself with language although it is not my proper task here. ["The Thought," 519 n; O 66n]

Frege's primary concern with knowledge and thought is also explicit throughout his career. It occurs in the title of his first great work: "*Concept-writing: A Formal Language, Modeled on that of Arithmetic, of Pure Thought*." And it recurs in his statements on his task as logician:

> It is possible for one sentence to give no more and no less information than another; and, for all the multiplicity of languages, mankind has a common stock of thoughts. . . . the task of logic can hardly be performed without trying to recognize the thought in its manifold guises. [*Translations*, 46 n; "Über Begriff und Gegenstand", O 196 n]

> Neither logic nor mathematics has the task of investigating souls and the contents of consciousness whose bearer is a single person. Perhaps their task could be represented rather as the investigation of the mind, of the mind not of minds. ["The Thought," 531; O 74]

Frege's interest was cognitive. He sought to understand the abstract structures and logical laws which were in his view the essence of thought and a basis for knowledge. His approach to this domain was deeply original in its opposition to psychologism and its lack of interest in scepticism. (In this

[6] *Begriffsschrift*, in *Frege and Gödel*, ed. J. van Heijenoort (Cambridge, Mass.: Harvard University Press, 1970), 70; *Begriffsschrift und andere Aufsätze*, ed. I. Angelelli (Hildesheim: Georg Olms, 1977), p. xii. Frege also disavows primary concern with language in *Translations*, 54 "Über Begriff und Gegenstand", O 204; 58 "Über Sinn und Bedeutung", O 27–28; and in *Nachgelassene Schriften*, 7; "Logik" (zwischen 1879 und 1891).

regard it is fruitful, if one does not push the point too far, to see him as an heir of Kant, divested of the trappings of psychology and the fear of scepticism, but continuing an investigation into the abstract structure of cognition.) Equally original were Frege's rigorous, mathematical methodology and his reliance on linguistic structure as the starting point for analysis. But analysis of language was only the means, or to borrow Wittgenstein's metaphor, the ladder, by which one arrived at an understanding of language-independent thought.[7]

II

"Meaning" is, as mentioned, a highly adaptive term. And perhaps one could harmlessly (and trivially) identify sense with "meaning" in a favored sense of "meaning," a sense that allows for the contextual promiscuity of *Sinn*. But in fact, discussions of Frege—especially in recent times—have not made this allowance.

The identification of sense with meaning began with Russell, Frege's earliest English commentator, who simply translated "Sinn" as "meaning."[8] Subsequent translations have mostly been more circumspect. But an important exception is Feigl's translation of "Über Sinn und Bedeutung." Although Feigl usually translates "Sinn" as "sense," there is a significant slip in the translation of the third sentence in passage (*B*), where "meaning" is exchanged for "sense," apparently as an equivalent translation of "Sinn": "Whoever accepts this sense [Plato's disciple and the teacher of Alexander the Great] will interpret the meaning of the statement 'Aristotle was born in Stagira' differently from one who interpreted the sense of 'Aristotle' as the Stagirite teacher of Alexander the Great."[9] This translation was probably influential, as we shall see.

The explication of *Sinn* in terms of meaning may be found in a multitude of sources. For example, Church—Frege's most powerful exponent—wrote in his review of Carnap's *Introduction to Semantics*:

> Frege makes this same distinction between the intensional meaning, the *sense* (Sinn) which a name expresses, and the extensional meaning, the *designatum* (Bedeutung) which the name denotes or designates. . . . Briefly, the sense of an expression is its

[7] See also *The Foundations of Arithmetic*, trans. J. L. Austin (Evanston, Ill.: Northwestern University Press, 1968), pp. iii–iv.; *Translations*, 59 "On Sense and Denotation"; O 29; *Nachgelassene Schriften*, 1–8; "Logik" (zwischen 1879 und 1891); 142–161; "Logik" (1897); "The Thought," 511n, 534; O 61, 76; and *passim*. For numerous explicit parallels between Kant and Frege, see *The Foundations of Arithmetic*. A broader discussion of the influence of Kant may be found in Hans D. Sluga's "Frege and the Rise of Analytic Philosophy," *Inquiry*, 18 (1975), 471–98, secs III and VII.

[8] Bertrand Russell, *The Principles of Mathematics* (New York: W. W. Norton & Co., 1902), 501.

[9] "On Sense and Nominatum," trans., H. Feigl, in H. Feigl and W. Sellars (eds.), *Readings in Philosophical Analysis* (New York: Appleton-Century-Crofts, 1949), 86. See G. Frege, *Kleine Schriften*, Angelelli ed. (Hildesheim: Georg Olms, 1976), 144; "Über Sinn und Bedeutung", O 26–27.

linguistic meaning, the meaning which is known to anyone familiar with the language and for which no knowledge of extra-linguistic fact is required; the sense is what we have grasped when we are said to *understand* the expression.

Carnap follows suit, writing, "The concepts of sense and intension refer to meaning in a strict sense, as that which is grasped when we understand an expression without knowing the facts."[10]

These explications are based on a passage in "On Sense and Reference":

(*C*) It is natural, now, to think of there being connected with a sign . . . besides that to which the sign refers, which may be called the reference of the sign, also what I should like to call the *sense* of the sign, wherein the mode of presentation is contained. . . . The sense of a proper name is grasped by everybody who is sufficiently familiar with the language or totality of designations [signs] to which it belongs. . . . ["On Sense and Reference," 57–58, O 27]

Later Frege writes, "A proper name . . . *expresses* its sense, *stands for* or *designates* its reference" (ibid. 61, O 31).

The association of sense with signs and the claim that understanding the language is sufficient for grasping the sense of a "proper name" certainly suggest the identity of sense and meaning. But the suggestion is misleading. Frege's remarks in (*A*) and (*B*) about the sense of context-dependent expressions undermine the identity. In fact, passage (*B*) cited above, is appended as a footnote to (*C*). This footnote should have served as warning against strictly identifying meaning (or what is "grasped" by everyone sufficiently competent in the language) and sense. For in the footnote, (*B*), Frege points out that proper names like "Aristotle" may have senses that are *not* "grasped" by everyone who is competent in the language.[11] Passages (*A*) and (*B*) indicate that Frege, in using the term "sense," was primarily concerned with mode of presentation, with the objective content of thoughts, rather than with the meaning of linguistic expressions. The objective thought content expressed with sentences containing proper names like "Aristotle" was regarded by Frege as (normally) publicly accessible, but not purely by virtue of mastering

[10] Alonzo Church, "Carnap's *Introduction to Semantics*," *The Philosophical Review*, 52 (1943), 301; Rudolf Carnap, *Meaning and Necessity* (Chicago: University of Chicago Press, 1964), 125. Although more elaborate in his discussion of Frege's notion of sense, Michael Dummett, in *Frege: Philosophy of Language*, continues in substantially the tradition that Church and Carnap established. See, for example, pp. 2, 92, 364, 584–585, 589, and *passim*.

[11] In interpreting passages (*B*) and (*C*), one must be careful with the term "proper name." Immediately before passage (*C*), Frege stipulates a special, technical and misleadingly broad use for the term: "by 'sign' and 'name' I have here understood any designation representing a proper name, which thus has as its reference a definite object (this word being taken in its widest range). . . . The designation of a single object can also consist of several words or other signs. For brevity, let every such designation be called a proper name." The term "proper name" in passage (*C*) should be taken in this special, broad sense—as applying to any singular term, including definite descriptions. Indeed, in (*C*) Frege seems to have in mind singular terms that might occur in a context-free, "perfect" language usable in a demonstrative science. In the appended footnote, passage (*B*), Frege focuses on proper names like "Aristotle"—proper names in the ordinary, narrow sense of the term. *These* proper names constitute an exception to the generalization articulated in (*C*).

the language. One would need to know something about the speaker and the context as well.

Language, for Frege, is the prime or only instrument for expressing objective language-independent thoughts. But the rules governing language interest him only insofar as they illuminate the structure of such thoughts. Frege would perhaps have granted that meaning and sense are identical in a "perfect," context-free language. But this would be because such a language would be perfectly fitted to express thought contents.

It must be said that Church and Carnap were interested only in such context-free languages. So although their explication fails to give the sense, or meaning, of "sense" ("Sinn"), it can perhaps be seen as at least extensionally accurate, given their purposes. But the explication carried the seeds of misunderstanding, seeds that bore fruit in the next generation's controversy over proper names.

Since Wittgenstein's *Investigations* numerous objections have been flung at Frege's theory (or remarks) about the sense of proper names. Several of these are undermined or seriously weakened by the misunderstanding we have been discussing. Perhaps the simplest is Searle's first objection:

> Do proper names have senses? Frege argues that they must have senses, for he asks, how else can identity statements be other than trivially analytic . . . [But, if Frege is right, in identities involving proper names like 'Tully = Cicero'] each name must have a different sense, which seems at first sight most implausible, for we do not ordinarily think of proper names as having sense at all in the way that predicates do; we do not, e.g., give definitions of proper names.[12]

The point that we do not give definitions to proper names is in effect acknowledged by Frege in (*B*). The sense or information value of proper names and other context-dependent devices differs from their meaning (if any) precisely because no linguistic rule or dictionary entry can capture the different information that might be carried by these devices in different contexts. Searle's objection could be considered relevant only if this difference were ignored.

This misunderstanding of Frege's view of names is accompanied by a related misunderstanding of his view of general terms. Kripke writes,

> Mill says that *all* 'general' names are connotative; such a predicate as 'human being' is defined as the conjunction of certain properties which give necessary and sufficient conditions for humanity—rationality, animality and certain physical features. The modern logical tradition, as represented by Frege and Russell, seems to hold that Mill

[12] John Searle, "Proper Names," *Mind*, 67 (1958), 166–173, repr. in *Readings in the Philosophy of Language*, Rosenberg and Travis, eds. (Englewood Cliffs, Prentice-Hall, NJ: 1971), 212. Page references will be to this volume. Saul Kripke, in *Naming and Necessity*, in D. Davidson and G. Harman (eds.), *Semantics of Natural Language* (Dordrecht: Reidel, 1972), 255, in effect gives the same argument.

was wrong about singular names, but right about general names. ["Naming and Necessity," 322]

Of course, Frege did not think that the sense of a nonindexical expression (which we will presume "human being" to be) varies with the speaker. The sense such an expression expresses is determined by the relevant facts—conventional, environmental—about the speaker's language. But there is not the slightest evidence that Frege thought that this sense was in general, or often, to be defined by "properties," or even senses, expressed in *other* terms of the language. Frege's test for whether two expressions, F and G, express the same sense is whether it is possible to believe the thought that ... F ... and fail to believe the thought that ... G. ... On this test "human being" and any expression made up of terms for rationality, animality and certain physical characteristics will clearly fail to have the same sense.

A second argument, which I shall call the rigid designator argument, can be found in Searle but is more thoroughly developed by Kripke. Searle writes (as if he were elaborating or extending Frege's theory):

> Suppose we agree to drop "Aristotle" and use, say "the teacher of Alexander", then it is a necessary truth that the man referred to is Alexander's teacher—but it is a contingent fact that Aristotle ever went into pedagogy....

Kripke writes

> A proper name, properly used, simply was a definite description abbreviated or disguised. Frege specifically said that such a description gave the sense of the name.

> Frege and Russell certainly seem to have the full-blown theory according to which a proper name is simply synonymous with the description which is used to replace it.

> If the name *means the same* as [a] description ... it will not be a rigid designator. It will not necessarily designate the same object in all possible worlds, since other objects might have had the given properties in other possible worlds, unless (of course) we happened to use essential properties in our description. So suppose we say, 'Aristotle is the greatest man who studied under Plato'. If we used that as a *definition*, the name 'Aristotle' is to mean 'the greatest man who studied under Plato'. Then of course in some other possible world that man might not have studied under Plato and some other man would have been Aristotle....[13]

This argument has been taken by some to reduce to absurdity Frege's view that proper names have a sense at all, let alone the sense of a definite description.

To begin with, it must be repeated that it is not Frege's view that proper names are synonymous with, have the same meaning as, or are abbreviations of definite descriptions. It is perhaps significant that in supporting his gloss of "sense" in terms of meaning, Kripke quotes the Feigl translation earlier criticized (*op. cit.*, 257).

[13] Searle, "Proper Names", 217; Kripke, *Naming and Necessity*, 255–257.

As an argument against the view that proper names have sense, the rigid designator argument has almost nothing to be said for it. (Kripke does not strictly present it as such, but many have taken it in this way.) Sense was introduced to account for cognitive content. Even if one supposed that the referents of names were all that mattered in analyzing their role in discourse bearing on necessity, one would have no reason at all to think that different names do not sometimes make different contributions to cognitive content. Indeed, the failure of substitution of co-referential names in belief contexts and Frege's paradox of identity—which were the primary phenomena to be explained in terms of sense—are altogether ignored by the rigid designator argument.

A more limited conclusion that has been drawn from the argument is that proper names do not (ever) have the sense of definite descriptions. (It is unclear whether Frege thought that the senses of proper names are always descriptive, but he did seem to think they sometimes are.) Even taken this way, the argument is unsound. But discussing it will require some detail.

The premise of the argument—that proper names are always rigid—has sometimes been disputed. It is said that proper names do not always function as rigid designators. It is replied that in such cases, we are not dealing with genuine proper names. It is counter-replied that the reply reduces the original claim to a stipulation about "proper name" and has no theoretical interest. This is denied. And so forth. In my view, talk of proper names as themselves being rigid is a mistake. Proper names are context-dependent referential expressions which are usually *used rigidly*, but which sometimes in certain anaphoric occurrences are used nonrigidly. Still, I think, Kripke and Searle had a genuine insight into the overwhelmingly normal use of proper names. This use often differs from that of definite descriptions. In discussing the argument, I shall simply grant its premise, at least to the extent of assuming that *for purposes of interpreting natural language discourse*, uses of proper names should always be treated as rigid.

As a first step in evaluating the argument, I want to compare Frege's approach to these matters with a more widely discussed approach, most naturally associated not with Frege, but with Russell. On the Russellian approach, one might attempt to accommodate the rigid designator argument, while maintaining that proper names have the sense of definite descriptions, by claiming that proper names *always* have wide scope. On this account, proper names' having sense is compatible with their referential rigidity, just as a definite description's having sense is compatible with its having wide scope. The idea can be illustrated as follows. "It is not necessary that Aristotle was a teacher" is taken to have something like the form:

(1) (ιx) Aristotle (x) [-Nec (Teacher (ιx) Aristotle (x))]

to use Russell's own notation untouched by his theory of descriptions. ("Aristotle is such that it is not necessary that he was a teacher.") Since the name is outside the scope of the necessity predicate or operator, its referent must be taken to be its actual referent. This holds even if the name is associated with a descriptive sense in the context. (I improve on (1) in note 15a below.)

Unlike this Russellian line, the Fregean approach does not rest on a claim about logical syntax. It makes a purely semantical point. On this approach one can accommodate (or express) the rigid designator intuition by saying that proper names always maintain their customary referent in counterfactual contexts. Such a view is compatible with holding that proper names are associated with descriptive senses in those contexts. (Actually, Frege wrote nothing about metaphysical modality. He might have taken any of various lines in response to Searle and Kripke's modal intuitions. I am concerned only to show that he was in a position to accept them.) Since the senses of proper names vary from person to person, a Fregean might reason, such names are pragmatically well suited to those intensional contexts of natural language where fixing on the referent is more important than conveying a particular way of thinking about the referent. Since in modal contexts (unlike belief contexts), there is no general reason why the sense associated with proper names by a given person might be particularly important, proper names tend to serve the purpose of fixing on a referent. The (near) rigidity of names and other indexical devices is thus the offspring of a marriage of convenience between cognitively promiscuous linguistic devices and contexts where *Sinn* does not matter.

Kripke is at pains to distinguish between the rigid–nonrigid distinction and the wide scope–narrow scope distinction:

> The facts that 'the teacher of Alexander' is capable of scope distinctions in modal contexts and that it is not a rigid designator are both illustrated when one observes that the teacher of Alexander might not have taught Alexander (and in such circumstances would not have been the teacher of Alexander). On the other hand, it is not true that Aristotle might not have been Aristotle. [*Naming and Necessity*, n. 25]

These remarks do, I think, make it evident that the notion of being rigid is not the same as the notion of having wide scope: a definite description could have wide scope, yet not be rigid. (This is perhaps all that Kripke wanted to establish at this point: the intent is unclear.) But the remarks do not serve, as some have taken them, as a defense of the rigid designator argument against the Russellian response that proper names are rigid because they always have wide scope. To serve that purpose, they would have to be supplemented by the claim that since names do not have narrow scope, they do not have wide scope either. I see no reason to accept this claim, or to think that it has any force against the view that names always (or normally) have wide scope.

Moreover, the claim has no carryover to the Fregean response that names always have customary reference in counterfactual contexts. The analogous claim against Frege would be that since proper names, unlike definite descriptions, do not have oblique reference in counterfactual contexts, they do not have customary reference in such contexts. Obviously this conclusion is untempting. In fact, it is self-defeating, since the assumption of the rigid designator argument is that proper names have a constant customary reference. So far nothing in the argument has touched the view that proper names have sense, descriptive or otherwise.

Passage (*B*) intimates that ordinary proper names and other indexical devices do not occur in thought, or in a language perfectly suited to thought. Such a language, however, should be expected to express the thoughts which Searle and Kripke express with rigidly used proper names. Let [S(*A*)] abbreviate a (nonmetalinguistic) canonical expression for the oblique denotation, or sense, of any expression *A*, and let "The *F*" express the sense contextually associated with "Aristotle." Then as a first approximation, we have the following:

(*i*) -Nec (S ('Teacher') (The *F*))

("It is not necessary *of* the *F* that he be the teacher of Alexander.") Thoughts expressed in ordinary language with proper names are always such that the customary reference of the name (and of the name's sense) is maintained in counterfactual contexts. (I improve on (i) in note 15a below.)

This is not the end of the matter, however. Kripke's remarks contain a distinction between saying of Aristotle that it is not necessary that he is the teacher of Alexander, and saying that in a given counterfactual situation Aristotle is not the teacher of Alexander.

> Not only is it true *of* the man Aristotle that he might not have gone into pedagogy; it is also true that we use the term 'Aristotle' in such a way that in thinking of a counterfactual situation in which Aristotle didn't go into any of the fields and do any of the achievements we commonly attribute to him, still we would say that was a situation in which *Aristotle* did not do these things. [*Naming and Necessity*, 279]

The idea here may be that the appeal to wide scope can account for the first locution but not the second. (Again the intent is not clear.)

David Kaplan has defended a more explicit version of the argument Kripke may have had in mind.[14] The strategy is this. Take the proposition expressed by an occurrence of "Aristotle was a philosopher." Got it? Call the proposition "Ari." Now carry this proposition over into a counterfactual circumstance in which *a*) whatever description you associated with "Aristotle"

[14] The argument has been given on numerous public occasions. It also appears in a circulated mimeograph, "Demonstratives," 1977. As written, Kaplan's argument applies only to demonstratives, though verbally he has tried it out on proper names. Frege would have to deal with both forms. Incidentally, Kaplan's piece is quite sensitive to the distinction between Fregean sense and meaning, but like Perry's (note 2) it tends to see Frege as an errant theorist of meaning or propositions. In conversation, Kaplan has agreed that the strategy involving "@," discussed below, circumvents the rigid designator argument, though as noted, his orientation on these matters is somewhat different.

(abbreviate this description by "The *F*") was satisfied not by Aristotle or any other philosopher, but by someone else, and *b*) Aristotle remains as involved in philosophy as he ever was. Kaplan claims that Ari is true in the counterfactual circumstance, whereas the proposition expressed by "The *F* is a philosopher" would be false. Thus "Aristotle" could not, on the relevant occasion, be associated with what "The *F*" normally expresses; and the corresponding sentences (on the relevant occasion) could not have expressed the same proposition. The point of the argument is to show that one need not rely on contexts like "It is necessary that..." to bring out the difference between names and definite descriptions.

This supplemental argument has more force against the Russellian "wide scope response" than against the Fregean response. The "wide scope response" depends on using an expression like "it is necessary that" as a syntactical pivot around which a proper name could swing into wide position. The supplemental argument attempts to remove the pivot. The original Fregean response, however, does not depend on a pivot, since its force is purely semantical. I think that the supplemental argument should not convert a Fregean living in *Sinn*. Two general points are open to him.

The first is that the supplemental argument, at least as presented by Kaplan, is not clearly relevant to Frege's position. Kaplan invites us to consider the proposition expressed by "Aristotle is *F*," then has us carry this proposition over into a counterfactual circumstance. We are led to conclude that the proposition has a different truth-value than Frege would say it has. But the matter is not so simple. In particular, the term "proposition" is a source of trouble. I want to dwell on this point for a while.

There is a tradition, stemming from Russell and Church, of using the term "proposition" in interpreting Frege's "Gedanke." I think this tradition no less misleading than that of using "meaning" to interpret "*Sinn*." In 1943 Church wrote, "The translation of Frege's '*Gedanke*' as 'proposition' is clearly justified by his explanation, 'nicht das subjektive Thun des Denkens, sondern dessen objektiven Inhalt, der fähig ist, gemeinsames Eigenthum von Vielen zu sein'."[15] Church's motive in proposing his translation is to avoid any reference to inner mental occurrences of individuals. But neither this motive nor the passage Church quotes justifies translating "*Gedanke*" as "proposition." "*Gedanke*" means "thought." There is nothing contradictory or even unordinary about a thought had or considered by many people. Frege himself argues that his use of "*Gedanke*" is in agreement with ordinary usage. He emphasizes that thought (*Gedanke*) is not the act of thinking (*Denktat*) and is not to be seen as produced by thinking. He continues:

[15] Alonzo Church, Review of Quine's "Notes on Existence and Necessity," *The Journal of Symbolic Logic*, 8 (1943), 47. The German runs: "not the subjective act of thinking, but its objective content, which can be the common property of many." The passage occurs in "On Sense and Reference," n. 7, O 32. See Russell, *The Principles of Mathematics*, 502 ff.

But my construal stands in agreement with many ordinary ways of speaking. Doesn't one say that the same thought is grasped by this person and that, that someone has repeatedly thought the same thought? Now if the thought arose only through thinking, or consisted in thinking, the same thought could arise, vanish, and arise again—which is absurd. [*Nachgelassene Schriften*, 147–149; "Logik" (1897)]

This passage, and others, clearly requires translating "*Gedanke*" as "thought." In using the term to interpret Frege, one must simply remember that it is thought contents that Frege intends, not occurrences in individuals' minds.

Like "meaning," "proposition" is vague and multi-purposed. But using it to translate "*Gedanke*" obscures the fact that Frege explicitly tied his investigations of logic to Kantian issues about (normative) laws of thought and judgment. Translating "*Gedanke*" as "proposition" tends to assimilate Frege's work to a tradition which, though heavily indebted to him, is more narrowly concentrated on issues in the semantics of natural language. Frege's intent is clear even where he is not using the term "*Gedanke*": "Neither logic nor mathematics has the task of investigating souls and the contents of consciousness whose bearer is a single person. *Perhaps their task could be represented rather as the investigation of the mind, of the mind not of minds*" ("The Thought," 531; O 74; italics mine.) If one understood "proposition" in terms of thought content, one would be on the right track. But typically, the notions of abstract thought and judgment, which were paramount for Frege (and nineteenth century logicians generally), are ignored in favor of the notion of linguistic meaning, which was expressly secondary. (See n. 6 above and "The Thought," nn. 1, 3, 5; O 6l n, 62 (second note), 69 n.)

Kaplan argues that the *proposition* expressed by an occurrence of "Aristotle was a philosopher" would be true even in the counterfactual circumstances in which the definite description that allegedly expresses the sense associated with the occurrence of "Aristotle" denotes some nonphilosopher. (Kaplan glosses "proposition" as "what is said" in the indirect discourse sense.)

The Fregean can reasonably reply that his interest is in thoughts, not "propositions." When a person uses "Aristotle was a philosopher," he thinks a thought and associates, contextually, a sense or thought component with the name. To isolate the thought expressed by the sentence, one must get at the cognitive significance of the name for the person on the occasion of use. One might *say of* Aristotle that he was a philosopher. We might even say that one thinks *of* Aristotle that he is a philosopher. (See (*i*).) But the *thought* expressed by an occurrence of "Aristotle was a philosopher," the Fregean might continue, involves the sense of a definite description. Nothing in the argument bears on whether the person's thought remains true or not in the counterfactual circumstance. For the Fregean, that depends on what thought component the person associates with "Aristotle."

The Fregean might concede intuitions Kaplan's argument plays upon. He might concede that Aristotle might have been a philosopher even as the *F* remained nonphilosophical (where "Aristotle" maintains customary reference). He might concede that the "proposition" that Aristotle was a philosopher would be true in such a circumstance. In fact, the Fregean can provide a reconstruction of the relevant notion of proposition. (Kaplan's technical term is "content.") The notion is gotten by taking the thought expressed by the relevant occurrence of "Aristotle was a philosopher" and replacing the sense expressed by "Aristotle" with its (actual) denotation. Holding the denotation constant, one considers whether it (he) would be a philosopher in the counterfactual circumstance. The result is an artificial construct called a "proposition." (See (*i*).) Such a proposition may indeed be counted true in the counterfactual circumstance in which the thought, or the proposition, that the *F* was a philosopher is not true. But the mere availability of such a notion of proposition does not at all count against carrying out an investigation of cognitive phenomena that appeals to thoughts. As long as such propositions can be constructed out of Fregean thoughts, the basic Fregean viewpoint remains intact. (See note 3.) I think this construction cannot in general be carried out. (See *opera*, note 15a below.) But nothing in the rigid designator argument shows this.

The Fregean might concede more: that in many cases a proposition in this sense is all that is really *communicated through ordinary language* when proper names, or other indexicals are used. ("So long as the reference remains the same, such variations of sense [in ordinary language] may be tolerated..." See (*B*).) Many modal contexts, indirect-discourse contexts, and even belief contexts in ordinary language ignore the possibly idiosyncratic senses associated with indexicals or proper names to concentrate on the publicly salient referent. On the other hand, these propositions do nothing by themselves to solve the problems of cognitive value and oblique belief contexts which thoughts were introduced to solve.

So far, I have argued that the Fregean can leave it open whether the *senses* contextually associated with names apply rigidly to their referents. But a second point is available to him. He can simply say that proper names are contextually associated with senses only of rigid descriptions: The sense of a name must apply to the same object under consideration of all counterfactual circumstances, the same object it applies to in nonmodal identity contexts. Nor is this much of a restriction. Frege could have said that "Aristotle" is contextually associated with the sense expressed by "the @ pupil of Plato who taught Alexander the Great," where "@" functions to insure that the sense of the ensuing description does not waffle or fail in its denotation as one considers different possible circumstances.[15a] The device

[15a] This device can be used to provide a simplification of the strategy represented by (*i*) above. The Fregean could replace (*i*) with: -Nec (S ('Teacher')S ('The @ *F*')). (See note 3.) This sort of definition

can be fitted to any description, and its availability by itself shows that the supplemented rigid designator argument does not cut very deeply as an objection to Frege.

None of the preceding is meant to imply that non-Fregean theories of cognitive content are impossible, or even that Fregean *Sinn* is justified. Indeed, I favor a non-Fregean theory. Rather it is to say that the rigid designator argument, in all its known forms, presents no counterexample to a Fregean theory of thoughts and does not touch the phenomena that the theory was introduced to explain.

A further objection typically raised against Frege is that sentences of the form "Aristotle was the *F*" (where "the *F*" represents a description that gives the sense of "Aristotle") are not *analytic*. It should be clear that on Frege's view no linguistic rules will determine that the utterance is true. So analyticity in its modern sense, like meaning, is not at issue. On the other hand, what is expressed by a particular utterance of the sentence would be (virtually) logically true, on Frege's view. Is this plausible?[16] To answer the question it will be important to discuss a pair of related objections that have been brought against Frege.

Many have noted that it is epistemically possible for the description a person associates with a name to turn out to be false of the name's object. Thus, we could discover that Aristotle wasn't a pupil of Plato or a teacher of Alexander. But suppose that a student, knowing little else to associate with "Aristotle," associates only these descriptions. Do we really want to say either that the student uttered a logical truth or that he failed to refer to

runs into trouble as applied to belief contexts, trouble I have discussed in "Belief *De Re*," sect. III, and "Kaplan, Quine and Suspended Belief," *Philosophical Studies*, 31 (1977), 197–203. But the trouble, which is also trouble for (*i*) understood in the Fregean way, stems from epistemic considerations and is not created by the rigid designator argument.

"@" differs from "actual" only in that "The @ *F*" picks out the actual *F* (in "our world") not only when it is used in an actual context and evaluated under counterfactual circumstances, but also if it were used in counterfactual circumstances. (There is no need for Frege to find an exactly corresponding locution for "@" in natural language.) This distinction between "actual" and "@" deprives one of the basis commonly used to argue that "actual" is an indexical. See David K. Lewis, "Anselm and Actuality," *Nous* 4 (1970), 175–88. It should be noted that Frege is *not* committed to cashing out "@" in terms of the senses of definite descriptions. As I mentioned earlier, he is nowhere committed to associating all singular expressions with the senses of definite descriptions. In fact, *obliquely occurring singular terms function as singular terms and appear to have nondescriptive senses in his system.* See his letter to Russell, 12/28/1902 in Gottlob Frege, *Wissenschaftlicher Briefwechsel*, ed. H. Hermes, F. Kambartel, and F. Kaulbach (Hamburg Felix Meiner, 1976).

Moreover, it is quite doubtful that "@" need be or contain a singular term. I shall not go into this matter in detail since it is complex. But in any discussion of Frege and modality, it should be remembered that Frege's semantical method as applied to modal contexts does not yield the model-theoretic or possible-world framework. Nor is there any simple argument that it should use such a framework.

[16] Searle, "Proper Names", 215; Kripke, *Naming and Necessity*, 255, 257–258. In calling sentences of the form F(The F) "virtually logically true" I am hedging an issue over how to interpret such sentences (or their occurrences) if the singular term fails to be uniquely satisfied by an object. In some logics, the sentence (occurrence) is nevertheless true; in others, it is not. In either case, the form shares the uninformativeness of logical truths.

Aristotle on the occasion when he said "Aristotle was the greatest pupil of Plato and the teacher of Alexander"? This rhetorical question is backed by a corollary claim that it is implausible to say that the pupil expressed a different proposition than, say, his teacher, who associated other and perhaps more fortunate descriptions with the name.[17]

We may begin with the corollary claim. There are certainly notions of proposition according to which it *is* implausible to differentiate between the pupil's and teacher's propositions. But "proposition" is not Frege's term. He explicates *Sinn* in terms of mode of presentation to a thinker and counts the sense expressed by a sentence a thought. It is considerably more plausible that in some sense or other, the pupil was thinking a somewhat different thought from the teacher on the relevant occasion. Of course, there is intuitively something the teacher and pupil said, believed, communicated in common. And it is a drawback of Frege's account that he does not address this point. But the sin is more clearly one of omission than of commission. It is evident that with proper names and other indexicals, Frege was more impressed by the individual's information than by mankind's "common stock of thoughts," which he highlighted in non-indexical constructions. It is clear that people with nonoverlapping descriptions associated with a name may not communicate very well. Their thought contents may be in principle publicly accessible, but in fact idiosyncratically entertained. The intuitions that Frege was trying to account for seem genuine. Whether or not his account is optimal, it is not a conclusive line of objection to point to other intuitions he ignored.

Is it plausible that the pupil expressed a logical truth or failed to refer, in the circumstance that we are imagining? Again, it seems intuitive that in some sense the student said and believed something false about Aristotle. And Frege says nothing about this intuition. On the other hand, it is not obvious, given the student's impoverished and mistaken information, that there is no sense in which the student's own thinking failed to pick out Aristotle. It seems to me that nothing much is to be gained by insisting on the point against Frege. There is no evident reason why the student's performance cannot be evaluated from a variety of viewpoints.

The question does arise whether Frege's theory is equipped to account for the intuitions he passed over. This is a complicated question that I shall not discuss in detail, since Frege said nothing about it. But a word is in order. It is certainly open to Frege to claim that while, in a sense, the student's own thinking failed to pick out Aristotle, we can also hold that the student made a statement and held a belief that was false of Aristotle because the student intended to refer to *whomever the teacher referred to*. The

[17] Wittgenstein, *Investigations*, § 79, trans. G. E. M. Anscombe (Oxford: Basil Blackwell, 1958), 79; Searle, "Proper Names", 218; Donnellan, "Proper Names and Identifying Descriptions", in Davidson and Harman (eds.), *Semantics of Natural Language*, 361–362.

teacher (and we) associate senses with the name (Frege might presume) that are more fortunate than the student's. A sense given by the italicized expression would, on these assumptions, pick out Aristotle. And the student would, from the viewpoint of this evaluation, have said and believed something false about him, while also expressing a thought that failed to pick out Aristotle at all. Alternatively, the student may be seen as using the proper name anaphorically to express the very sense that the teacher associates with the name. Thus the sense *expressed* by the student would be different from, and additional to, the sense that he grasps in his own thinking. On either view, the intuition that the student said and believed something in common with the teacher could then be explained in terms of their saying and believing the same thing *of* the same person. (This is a locution Frege himself used. See *Nachgelassene Schriften*, 275; "Aufzeichnungen für Ludwig Darmstaedter" (1919).) I do not present these remarks as an adequate solution. I do not think that they are. What I want to indicate is that if one recognizes the variety of intuitions at issue, it is not clear—*on the basis of considerations raised so far*—that Frege's theory cannot do reasonable justice to them. It is a peculiarly philosophical mistake to assume a person's linguistic performance or propositional attitude is to be evaluated from only one viewpoint. The mistake is common to many of Frege's critics. Frege may have made it too.

These remarks can be extended to apply to the question of whether it is plausible that an ordinarily informed person might express a logical truth by saying "Aristotle is the pupil of Plato and teacher of Alexander." (We shall suppose here that in fact Aristotle *was* the pupil of Plato and teacher of Alexander.) Frege seems to have anticipated this issue in passage (*B*)—picking an example in which a sentence expresses a virtual logical truth in one mouth and an ordinary factual assertion in another. He would say that a sentence of the form "Aristotle is *F*" can be judged not to express a logical truth only from a viewpoint in which the sense attached to "Aristotle" is independent of that expressed by "*F*." Persons using different senses for "Aristotle" connect themselves to one another by guessing each other's senses (near enough) and perhaps by carrying an auxiliary sense such as "whomever the rest of them are referring to." The picture is not as neat or precise as one might like. But it is not impossibly unintuitive.

We now turn to a brace of objections to Frege that seem more serious. The first is that Frege gave no means of determining the sense of a name used by a person. Except in a few cases, the person is likely to be at a loss to say what the sense of a name is. If we allow all the descriptions at his disposal to count, then (assuming he is well-informed) his sense will be too vulnerable to reference failure even to sympathetic intuitions. Moreover, if such failure did not occur, all singular thoughts using the name would be virtual logical truths. On the other hand, it is hard to see how to delimit the sense, if the person himself is not able to do so.[18]

Frege's position on how senses are determined is more complex than is commonly supposed. An attentive reading of passage (*A*) and a related passage in "The Thought" indicates that Frege believed that the sense associated with indexical expressions, in a context, is not determined by the speaker's beliefs or psychological state in any simple sense:

(*A*) If a time indication is made in present tense, one must know when the sentence was uttered to grasp the thought correctly. Thus the time of utterance is part of the expression of the thought . . . Although the thought [expressed by "Today . . . " and by "Yesterday . . . " in a given pair of contexts] is the same, the verbal expression must be different, to compensate for the change of sense which would otherwise be brought about by the different time of utterance.

(*D*) The words 'this tree is covered with green leaves' are not sufficient by themselves for the utterance; the time of utterance is involved as well. Without the time indication this gives we have no complete thought, i.e. no thought at all. ["The Thought," 533; O 76]

The sense appears to be determined by the context (by the time in these cases) in a way that does not completely depend on attitudes of the speaker.[19] The picture is not that of a person's assigning a sense to his indexical construction, but of his using an indexical construction and a sense's being assigned to the construction, and a thought to the sentence, by the context. Frege apparently does not assume that a person must be able to give a nonindexical account of the thoughts he expresses in indexical terms. He does leave some latitude for the speaker's intentions even in the use of ordinary indexical constructions. He suggests that someone could use "I" either in a "special and original" ("besonderen und ursprünglichen") way, or in the sense of "he who is speaking to you at this moment" ("The Thought," 519; O 66). But apparently the context plays a primary role in assigning the sense or thought component of indexical occurrences, independently of the speaker's other thoughts or beliefs.

This picture is subtle and intriguing. But as it stands, it presents only a direction—one very different from that which is commonly attributed—not a theory. It should also be noted that Frege does not appear to have applied the picture to proper names. (See passage (*B*).) Thus it remains unclear precisely how the senses of proper names or indexical expressions are contextually

[18] Wittgenstein, *Investigations*, § 79; Searle, "Proper Names" 214–215; Kripke, *Naming and Necessity*, p. 257; Donnellan, "Proper Names and Identifying Descriptions", sec. V.

[19] Hilary Putnam, in "The Meaning of 'Meaning' ", in *Philosophical Papers*, ii (Cambridge: Cambridge University Press, 1975), 218, in effect attributes to Frege the view that grasping meaning (or sense) is just a matter of being in a psychological state. I know of no place where Frege states or implies this view. Although he does not explicitly take issue with it, passages (*A*) and (*D′*), below, tend to controvert it.

determined. With names we rely more on the person; with other indexicals, more on his context. But a clear account of what senses are assigned, and how, is missing in both cases. Although serious, this objection is hardly knockdown. Taken by itself, it appears to demand merely that an incomplete story be completed or continued. The problem arises when one tries to continue it plausibly, a problem shaped by a final objection.

Frege required that senses be sufficiently complete to determine their associated referents by their very nature. The truth-value of a thought expressed by a sentence utterance of the form "Aristotle is ϕ" or "It is raining" must be determined purely by the eternal nature of the thought. The sense of a proper name or indexical construction is a timeless abstract entity that bears its relation to its denotation or referent in an eternal, context-free manner. It bears this relation in complete detachment from anything a person does or anything that happens to him. Senses are or are components of thoughts, but they determine their referents in regal independence of the thinker's activity. They are simply there to be grasped ("The Thought," 530–531, 533–534; O 73–74, 75–76; and *Nachgelassene Schriften*, pp. 147–149; "Logik" (1897)).

This conception leads to several fundamental difficulties with Frege's system. But the one I shall discuss here is special to context-dependent reference. The problem is that it seems intuitively implausible that a person who uses proper names and indexical constructions always has or grasps abstract thought components (I shall call them "concepts") that are sufficiently complete to determine uniquely and in a context-free way the things he refers to. For example, one might grant Frege the ploy of interpersonal cross-reference ("whomever he referred to") with proper names discussed earlier. But there may not always be a "back-up" person so readily at hand. And sometimes no person or group of persons meeting the condition of conceptual completeness will be available in one's community. There is no guarantee that such chains of cross-reference will always eventuate in someone with a complete sense or thought component.

The problem is even more stark with demonstratives. Frege writes

(D') But are there not thoughts which are true today but false in six months time? The thought, for example, that the tree there is covered with green leaves, will surely be false in six months time? No, for it is not the same thought at all. The words 'this tree is covered with green leaves' are not sufficient by themselves for the utterance; the time of utterance is involved as well. Without the time indication this gives we have no complete thought, i.e. no thought at all. Only a sentence supplemented by a time indication and complete in every respect expresses a thought. But this, if it is true, is true not only today or tomorrow but timelessly. [See *Nachgelassene Schriften*, 4–5; "Logik" (zwischen 1879 and 1891); "The Thought," 533; O 76.]

But it is not intuitively plausible that a person's conceptual abilities will always include a grasp of a context-free, complete "time indication." The problem is not just that he may not know what time it is. It is that an inventory of thoughts he believes, of the conceptual resources that enter his beliefs, may be insufficient to uniquely determine the time (or other object) to which he succeeds in referring indexically.[20]

Frege appears to be caught between two objectives that he had for his notion of sense. He wanted the notion to function as conceptual representation for a thinker (though in principle accessible to various thinkers, except in special cases). And he wanted it uniquely to determine the referents of linguistic expressions and mental or linguistic acts.[21] The problem is that people have thoughts about individuals that they do not individuate conceptually.

It is unclear how Frege would have dealt with this problem had it been forced upon his attention. I shall consider two sorts of responses. One preserves most of Frege's doctrine (though it is not compatible with everything he said), but seems implausible. The other gives up a fundamental tenet of the doctrine.

One response is to postulate *special* senses for proper names and demonstratives (as used in a context), senses that are not expressed by any other expressions in the language. These senses or concepts would completely fix their referents and would do so in a context-independent way, although they would be expressed or thought only contextually. One might say, for example, that the sense or concept expressed by "Aristotle," in a context, is that of being Aristotle. The sense of a use of "this" or "now" would perhaps be that of being *this* or being *now*.[22] One might see these senses or concepts as assigned by the context in the manner suggested above, so that they could not fail to have the "right" referent.

[20] For elaboration of this sort of point, though often with a focus on meaning rather than sense, see Wittgenstein, *Investigations*, *passim*; Strawson, *Individuals* (Garden City, NY: Anchor, 1963), 6–9; Donnellan, "Proper Names and Identifying Descriptions", sec. V, VII–IX; Kripke, *Naming and Necessity*, pp. 291–92; "Belief *De Re*," sec. IV. Even when a person *can* uniquely describe the objects he indexically refers to, the indexically expressed attitudes are intuitively not the same as their eternally expressed counterparts. For the former essentially involve the subject's *application* of attitudinal contents in a particular context to contextually identified entities. The latter need involve no such ability to make a contextually appropriate application.

[21] Frege, "The Thought," 533–534; O 75–76; "On Sense and Reference," 58; O 27–28. This function of determining the reference has also been misunderstood or distorted. I have elsewhere warned against identifying it with Kripke's pragmatic notion of fixing a referent. See Kripke, *Naming and Necessity*, 274–278; "Belief *De Re*," 356–357. There is also a widespread tendency to "operationalize" Frege's function of determining the reference—identifying it, say, with a method or procedure of verification. This is anachronistic, though perhaps not uninteresting from an historical point of view. Frege's notion is vaguer. Although some of his examples suggest this interpretation (e.g. the telescope analogy, "On Sense and Reference," 60; O 30), Frege simply requires that for every sense there be at most one referent. I should note that my use of "concept" in the text is to be strictly distinguished from Frege's "concept" (*Begriff*). Frege's term can be glossed as "a function whose values are truth-values"; first-level concepts are the referents of predicates. I use "concept" to apply to components of thoughts, in Frege's sense of "thought," which are not themselves thoughts—senses of expressions other than sentences. I shall broaden this use of "concept" in a few pages.

[22] This sort of move may be seen in variations on Carnap's idea of assigning individual concepts to proper names. See *Meaning and Necessity*. Of course, such concepts as Carnap conceived them are not sufficiently fine to serve Frege's epistemic purposes. It may also be seen as kin to the idea attributed to

I think that anyone not already bent on preserving a philosophical outlook will find this sort of response thin and implausible. It amounts to describing senses or concepts with all the theoretically required features without doing anything to assuage the original doubt that there are such senses or concepts. It is not clear what one is being told when it is said that the sense of "Aristotle" is the concept of being Aristotle. The expression "the concept of being Aristotle" does not suffice to convey what is intended, for it is just as context-dependent as the proper name. This insufficiency takes two forms. In the first place, there are lots of Aristotles—which one is intended? We seem to rely on the context to pick out the "right" one. But intuitively we do not—at least not always—rely on some contextually associated complete sense or concept which eternally determines the referent. In the second place, the name "Aristotle" may carry—even for a given Aristotle, a given speaker or thinker, and a given time—different cognitive values. "Aristotle is Aristotle," as used at a given time, may express a surprising discovery rather than a triviality (See my "Belief *De Re*", n. 2).

Analogous points hold for demonstratives. One can believe "what is expressed" by "*i* is *F*" (where "*i*" represents any demonstrative and *F*, a nonindexical predicate) in one context and disbelieve "what is expressed" by the same sentence in another context even though *a*) "*i*" applies to the same entity and *b*) intuitively one is not changing one's mind. For example, one might correctly believe that today is Thursday, and later in the same day, having thought twenty-four hours had passed, disbelieve that today is Thursday. Frege will need indefinitely many complete senses, even holding both the referent of the indexical and the indexical itself fixed. As before, the appeal to concepts or senses "complete in every way," that by their nature uniquely fix the referents of these context-dependent expressions, has no intuitive substance.

A side issue here, but one important to Frege, is that the appeal to special senses inexpressible in other terms provides a dubious basis for an account of communication. For given that those senses must be distinguished so finely (roughly to match the person's particular epistemic viewpoint in a context), and given that they cannot be explicated in nonindexical terms, it is difficult to see how they can ever be communicated. For each person's epistemic viewpoint, even in a given situation, is different. To be sure, Frege thought (plausibly) that communication with proper names and other indexicals is in

Boethius by Plantinga (and favored by Plantinga himself) of letting the proper name "Aristotle" express the property of being Aristotle. See Alvin Plantinga, "The Boethian Compromise," *American Philosophical Quarterly*, 15 (1978), 129–138. Here again, unless properties are distinguished as finely as senses expressed in a context, the idea will not suffice to meet Frege's epistemic demands—those that issue from the various forms of the paradox of identity. Versions of this general approach are also defended in Diana Ackermann, "Proper Names, Propositional Attitudes and Non-Descriptive Connotations," *Philosophical Studies* 35 (1979), 55–69; and Castañeda, "On the Philosophical Foundations", and "Perception, Belief, and the Structure of Physical Objects and Consciousness," *Synthese*, 35 (1977), 285–351. The approach-type has been proposed informally by several others.

some respects less reliable than with other expressions. And he held that each person has a sense for "I" that is *in principle* incommunicable to anyone else ("The Thought," 519; O 66). But Frege's point about "I" seems to have been a rather clumsy attempt to capture the mundane, but special fact that only *a* can take the first-person viewpoint toward *a*. And it is evident that Frege did not hold that all senses for indexicals are in principle incommunicable.[23] He gives examples of indexicals which he thinks can be communicated ("I" in one sense, "he," "you," "this," present tense, and proper names, "The Thought," pp. 517, 519, 533; O 64–65, 66, 76). But if senses or concepts are individuated finely enough to bear Frege's epistemic load (accounting for possible differences of belief), yet are counted rich enough to meet his semantical requirement (specifying by their very abstract natures a unique referent), it is hard to see how communication with indexical constructions could depend on them in practice, whatever one said about the matter in principle. It is not clear, however, what Frege thought about communication involving indexical sentences.

One might think of an indexical like "this" as contextually associated with a largely qualitative sense, something like what is presented in the visual field. (See "On Sense and Reference," p. 60; O 30.) But this idea increases intelligibility without achieving adequacy. It is in principle possible for identical visual fields to be associated with different referents. And no supplementation of the visual field with concepts will always be sufficient to make it plausible that the referent is completely determined by the nature of complete senses or concepts (as opposed to indexicals) that the person employs. (See n. 20.) The appeal to visual fields or qualitative senses does little to capture uses of "now" and tenses, or the first-person pronoun, or blind pointings, or indexicals based on memory (where imagery fades), or proper names of people or entities with whom we are not acquainted.

The troublemaker underlying these contortions is the assumption that a person using indexical constructions always thinks thoughts that are context-free and "complete in every way"—that the contents of a person's beliefs and so forth, are always completely conceptualized. What I mean to convey by "completely conceptualized" is that the truth-value of the cognitive content—of what the thinker grasps, thinks, or believes—is eternally fixed given its nature, given the kind of content it is. A trademark of a sense or Fregean thought component is that it can in principle be expressed on indefinitely many occasions. For nothing in its expression or in its being thought affects

[23] Perry, "Frege on Demonstratives", sees the appeal to incommunicability in the case of "I" as the result of pressure from some of the problems that we have discussed. I do not find this plausible. Appeal to incommunicability emerges only with "I," but the problems of incompleteness arise with other indexicals as well. These problems raise worries about how communication might occur in practice; but it is not clear why they should lead one to appeal to incommunicability in principle. There is also no clear evidence that Frege considered the problems we are discussing.

its referential relations. (See "The Thought", 530–531; O 73–74.) Its relation to its referent(s) is atemporal and depends purely on its own nature and the inventory of the world. The problem we have been discussing is that thoughts sometimes appear to be irreducibly context-bound. The response in terms of special senses tries to confine the context-dependence to the *expression* of thoughts, maintaining that the thoughts themselves are eternally self-sufficient. This is a consistent position. But it is intuitively implausible. Its implausibility emerges in the strained, ad hoc and inarticulate character of attempts to specify or evoke the senses or concepts that are supposed to be indexically expressed. There seems no natural means of transcending the indexical character of our thought expressions or thought ascriptions.

An entirely different response to these problems concedes the context-bound character of thoughts. Frege himself held views that could perhaps have been mobilized to yield such a response. In *The Foundations of Arithmetic* he appealed to Kant's notion of intuition in his discussion of geometry:

> A geometrical point, regarded by itself, cannot be distinguished from any other; the same holds for lines and planes. Only when more points, lines, planes are simultaneously apprehended in an intuition, does one distinguish them. When in Geometry general propositions are extracted from intuition, this can be explicated by the fact that the intuited points, lines, planes are not at all special (besondere) and thus can count as representatives of their kind.[24]

The notion of intuition, though somewhat vague, is fitted to the problem that we have raised for Frege. It is used to account for a thinker's apprehension of an entity without his being able to distinguish it conceptually from all other entities. The apprehension of it depends essentially on contextual, nonconceptual relations to it. Different contexts necessarily mark different intuitions.

The notion of intuition is treacherous. The task of explicating it tends to bring out the worst in philosophers: appeals to special kinds of infallible knowledge, which nevertheless cannot be conveyed, and the like. Indeed, the notion has tended to take on many of the objectionable features of the "special senses" discussed earlier. In my view, it is best to see the notion simply as indicating a thinker's contextual indexical application of concepts (constant nonindexical thought components) to individuals. Such application may be (perhaps always will be) backed or guided by further images, descriptions, concepts or the like. But these need not be sufficient to individuate the entities about which the thinker is thinking, nor need they be regarded as constituents of the intuition or of the relevant thought. A thought expressed in the form "that G is F" may be seen as an indexical, intentional application of

[24] Frege, *The Foundations of Arithmetic*, sec. 13; my translation here. I have tried to avoid using English technical, philosophical terms in translation unless they clearly match Frege's notions. Thus I translate "besondere" as "special" rather than as "particular." On the notion of intuition, see Kant, *Critique of Pure Reason*, A19/B33.

the concept *F* (in the just stated sense of "concept") to an entity which is described, but perhaps not completely individuated by the concept *G*. The context-bound application is a part of the representational function of the thought. There is nothing in the *nature* of the intuition, regarded apart from its context, that determines that it picks out the individual that it does. Individuative reference—even in thought—is partly a matter of the context in which concepts are applied by the thinker.

The view makes communication intelligible. A hearer can note the contexts in which indexicals are applied (rather than having to divine an idiosyncratically associated concept) and can utilize his notes to find the intended referent. A thorough discussion of the approach is out of place here. Suffice it to say that I think an appeal to some nonconceptual, context-dependent notion like intuition—or intentional, contextual application—is exactly what Frege needed to handle demonstratives, tense, and proper names.

Why did Frege not introduce this sort of notion? Throughout his career, his genius was focused on liberating logical theory from the vagaries of traditional epistemology. Understandably, he was deeply impressed with the explanatory power of his logical principles that reference (including truth-value) is a function of sense, and that the sense of a complex expression is a function of the senses of the parts. When he came to treating proper names and demonstratives, the impulse to apply these principles was overwhelming. Yet the notion of sense was originally introduced to deal with problems about informativeness—problems about knowledge and belief. And Frege was systematic enough to want his logical principles to do epistemological work in his discussion of context-dependent referential devices. The fundamental error of Frege's theory of these devices is that the logical principles, applied in their full strength, are epistemically implausible. It is not the case that for every indexical construction in each context in which it is successfully used, the user or thinker grasps a sense or concept that is complete enough by its very nature and apart from contextual application to specify the referent uniquely. Frege's focus on logic blurred his vision of epistemology. A similar short-sightedness has been inherited by most of Frege's critics, interpreters and followers. Influenced by his revolutionary approach to language and logic, they have tended to underestimate the depth of his insights and ambitions in epistemology. Indeed, recent thinking, far more than Frege, has ignored the demands of epistemology for the enticements of semantics.

There is a moral to be drawn here about criticisms of Frege: Only epistemically oriented criticisms are likely to be relevant to Fregean *Sinn*. Frege's notion marks a set of problems in the theory of cognition that cannot be reasonably ignored. Better acquaintance with *Sinn* should yield a richer understanding of the range of problems Frege bequeathed us. And in the light of such understanding, Frege's *Sinn* can be forgiven, if not justified.

Postscript to "Sinning Against Frege" (2003)

I would like to make one comment on what I wrote in this article. On pp. 223–230 I offered a response to the rigid-designator criticism of Frege's notion of sense. In a careful discussion of this issue, without specific reference to "Sinning Against Frege", Scott Soames argues against a defense of Frege that appeals to an actuality operator on definite descriptions.[1] "The actual F" used by someone in non-actual circumstances will pick out whatever satisfies the description in those non-actual circumstances, whereas a rigid proper name in the same circumstances, used by the same person, will pick out the bearer of the name in the actual world. So the two thoughts are different and the actuality-operator account has not provided a satisfactory account of the modal behavior of proper names in thoughts.

Soames points out that the modal behavior of proper names in uses within non-actual situations can be mimicked by a definite description if the definite description is fitted out with David Kaplan's *dthat*-operator.[2] He maintains that a descriptivist cannot accept this expedient, however, since *dthat*-operators in effect obliterate the descriptive content of the definite description. This would compromise the point of appealing to descriptions as senses or cognitive values.

None of this discussion touches the defense of Frege that I give against the rigid-designator objection in "Sinning Against Frege". The @ operator with which I outfitted definite descriptions is explicitly contrasted with the actuality operator in just the way that Soames takes intuitive usage and the *dthat*-operator to contrast with the actuality operator. It denotes the actual referent of a definite description, no matter whether that description is used in the actual circumstances or in counterfactual circumstances. It anticipates and specifically deals with exactly the problem that Soames raises.[3] The @ operator that I introduced is, however, not explained in such a way as to obliterate content, as the *dthat*-operator is. It retains the descriptive content in the content of the singular term.

So how are we to understand uses of such descriptions (as stand-ins for proper names) in counterfactual circumstances? According to the descriptivist Fregean, if the name is used in the counterfactual circumstance with the same sense that it expresses in the actual circumstances, it will be used in the counterfactual circumstance as expressing the same descriptive content that it is used to express in the actual circumstance; and (given the rigidifying

[1] Scott Soames, "The Modal Argument: Wide Scope and Rigidified Descriptions", *Nous*, 32 (1998), 1–22.

[2] Ibid. 17.

[3] Cf. note 15a above.

operator) the descriptive content will, in the counterfactual circumstance, determine the same referent that the name denotes, and the expressed descriptive content determines, in the actual circumstance.

Of course, Frege showed little interest in modality. And it is open to him to say that in counterfactual circumstances there is no general reason to believe that a name will be used to express the same sense as it does in a given actual circumstance. Still, the @ operator gives the descriptivist Fregean the wherewithal to say what it would be to express the same sense in both actual and counterfactual circumstances—while maintaining a rigid reference. And this device enables the descriptivist Fregean to mimic correctly the modal behavior of rigid proper names in his theory.

So Soames's defense of the rigid-designator criticism does not come to grips with the reply that I gave. Soames does not claim that it does. But others have maintained that it undermines my discussion. I continue to believe that the rigid-designator criticism is not decisive against a descriptivist Fregean view.

On the other hand, I want to emphasize that my defense of a descriptivist Fregean view against the rigid-designator criticism has a strictly "for the sake of argument" character. In the first place, I do not think that Frege was committed to anything as general as descriptivism. More importantly, I do not accept any version of the Fregean position—much less any generalized descriptivism. I think that many proper names (or uses of them) do not express (even in the most liberal sense of "express") thought components that determine their referents independently of any occurrences in time. And I think that there are thought components that are in no sense descriptive. In fact, the thought components normally expressed by proper names have no *descriptive* elements at all in their representational contents.[4]

[4] These non-descriptive elements are not only the applications of demonstrative elements in thought and the constant demonstrative elements themselves. They are also the predicative element in proper names. "Is a Smith" does not describe anyone. "Is a Smith" does not mean "is named 'Smith' ", although such phrases are sometimes approximately equivalent in extension. The metaphrase does not describe anyone either. For the background of these remarks, see "Reference and Proper Names", *The Journal of Philosophy*, 70 (1973), 425–439, and "Belief *De Re*", *The Journal of Philosophy*, 74 (1977), 338–362.

6 *Frege on Sense and Linguistic Meaning (1990)*

Frege's conception of sense has fathered all the major approaches to 'meaning' that have preoccupied philosophers in our century. Some of the progeny have developed the Fregean conception. Others have rebelled against it. All have borne its mark. Certain key elements in Frege's conception have, however, disappeared in interpretations of his view. These elements derive from his rationalist predilections and from his primary concern with idealized thought. The philosophical traditions most heavily influenced by Frege have been empiricist, preoccupied with language and suspicious of the notion of thought—idealized or not. So the neglect is natural. I think that some of the neglected elements in Frege's conception are of profound philosophical importance. They are critical to a philosophical understanding of language, cognitive processes, and conceptual change. I shall not, however, be building on Frege's conception in this paper. Rather I shall approach it historically.

I

Frege introduces the notion of sense by giving it three functions.[1] The first is that of accounting for 'cognitive value'. Senses are 'modes of presentation':

The following works by Frege are cited in the text by the abbreviations that follow their titles: *The Basic Laws of Arithmetic* (*BL*), trans. and ed. M. Furth (Berkeley: University of California Press, 1967); *Begriffsschrift und andere Aufsätze* (*B*), ed. I Angelelli (Hildesheim: Georg Olms, 1964; 2nd edn. 1977); *Foundations of Arithmetic* (*FA*), trans. J. L. Austin (Evanston, Ill.: Northwestern University Press, 1968; Oxford: Basil Blackwell, 1980); *Die Grundgesetze der Arithmetik* (*GG*) (Hildesheim: Georg Olms, 1962); *Kleine Schriften* (*KS*), ed. I. Angelelli (Hildesheim: Georg Olms, 1967); *Logical Investigations* (*LI*), ed. P. Geach (New Haven: Yale University Press, 1977); *Nachgelassene Schriften* (*NS*), ed. H. Hermes, F. Kambartel, and F. Kaulbach (Hamburg: Felix Meiner, 1968; 2nd edn. 1983); *Philosophical and Mathematical Correspondence* (*PMC*), trans. B. McGuinness and H. Kaal (Chicago: University of Chicago Press, 1980); *Posthumous Writings* (*PW*), ed. H. Hermes, F. Kambartel, and F. Kaulbach, trans. P. Long and R. White (Chicago: University of Chicago Press, 1979); *Translations from the Philosophical Writings of Gottlob Frege* (G & B), trans. and ed. P. Geach and M. Black (Oxford: Basil Blackwell, 1966); *Wissenschaftlicher Briefwechsel* (WB), ed. G. Gabriel, H. Hermes, F. Kambartel, C. Thiel, and A. Veraart (Hamburg: Felix Meiner, 1976). The sign "O" marks the pagination of the original publication of the cited article by Frege.

[1] These functions are developed throughout Frege's work after 1891. But they appear in succession in 'On Sense and Reference', 57–8, 58, 58–9, in G & B, O 26–27, 27, 28. *The Journal of Philo-*

ways things are presented to a thinker—or ways a thinker conceives of or otherwise represents entities in those cases where there are no entities. Not all modes of presentations are senses. But where modes of presentation are senses, they are associated with linguistic expressions. Being a sense is not essential to the entities that are senses. Being a (possible) mode of presentation to a thinker is what is fundamental. A sense is a mode of presentation that is 'grasped' by those 'sufficiently familiar' with the language to which an expression belongs. The second function of the notion of sense in Frege's theory is that of fixing the *Bedeutung*, the denotation or fundamental semantical value, of semantically relevant expressions. The third is that of serving as the denotation of expressions in oblique contexts.

Our primary concern is with the first function. Although the third is essential to Frege's overall theory, I think it less fundamental for understanding his conception of sense, at least initially, than the first two. The second function has been widely discussed, and I shall say only a few words about it now.

Frege explicates the notion of fixing a *Bedeutung* in a purely logical way: for each sense there is at most one *Bedeutung*. It is also clear, partly from the first function, that a sense is a way of thinking of *Bedeutung*. Beyond the foregoing, Frege says little.

The urge to say more has led some to interpret this second function in verificationist terms: sense is construed as a procedure for determining *Bedeutung*.[2] Thus the sense of a name would be a means of finding or recognizing an object; that of a predicate would be a way of determining whether an object satisfies it; that of a sentence would be a method of verification or falsification. To deal with a variety of problems of interpretation, these procedures have been regarded as highly idealized—for example, as what a collection of experts or a superior being would do. I think that there

sophy, 74. They are also discussed in my 'Belief *De Re*', (1977), sec. 4. References to Frege's writings will occur in the text, using abbreviations set out in the preliminary note. Where English and German paginations differ, both will be cited, separated by a slash. I am responsible for all translations of quotations from Frege, although often there will be little or no difference from already published translations. I will not say much on the vexed question of the translation of '*Bedeutung*'. I think that the current vogue of translating the term as 'meaning' is unfortunate, despite the support from ordinary translation practice outside of philosophy. The original choice, 'reference', was also unfortunate—though less so, in my opinion. The translation I prefer and have used here is 'denotation'. The translation goes back to Church, I believe. The advantage of this choice is that the term has an ordinary meaning ('meaning') which is roughly the same as the ordinary meaning of the term '*Bedeutung*' but also has various technical meanings that are not to be presumed the same as the ordinary meaning. In English-speaking philosophy, these technical meanings bear comparison to Frege's technical meaning for '*Bedeutung*', although the analogies (especially to Russell's and Mill's technical usage of 'denotation') must be handled with extreme caution. An alternative choice would be to leave the term '*Bedeutung*' untranslated. In the light of the moral of the present paper, as applied to '*Sinn*', this choice has much to be said for it: much gets lost and distorted under translation—especially translation between philosophical cultures.

2 Michael Dummett, *Frege: Philosophy of Language* (London: Duckworth, 1973), 488 ff.

is little justification in the texts for this procedural interpretation of Frege's notion of sense. I shall return to it later.

The main import of the second function lies in its connecting thought and judgement with truth, the *Bedeutung* of sentences. The first function indicates that the notion of sense is designed for an account of thought and judgement. The primary logically relevant function of thought and judgement is to 'strive after truth'.[3] Thus sense provides a logically relevant connection between judgement and truth. This connection grounds Frege's celebrated insistence that in construing the sense of an expression one consider only the expression's contribution to the deductive inferential potential and 'truth-conditions' of the sentences in which it is contained.

Let us now concentrate on the first function of the notion of sense in Frege's theory. The first function is to represent the way entities are presented to a thinker, or the way a thinker conceives of or otherwise represents entities in those cases where there are none. Frege's development of his theoretical notion of sense, through consideration of linguistic phenomena such as the paradox of identity and through his remark that sense is what is grasped by those 'sufficiently familiar' with a language, has led many to assimilate his conception of sense to modern notions of conventional linguistic meaning.[4] This interpretation cannot be sustained. My primary purpose is to explain in some detail why.

[3] I have discussed at some length the role of judgement and truth in Frege's system in 'Frege on Truth' (1986) (Ch. 3 above).

[4] This view was almost universally held until a few years ago. It has received its most comprehensive articulation in writings by Michael Dummett. Dummett's characterization is very complex and not to be caricatured. But despite its many strengths, it seems to me fundamentally a misrepresentation.

The notion of conventional linguistic meaning has itself received numerous characterizations. I shall not attempt another. I shall assume, however, that conventional linguistic meaning is what is understood by an ordinary speaker of a language, or by the 'most competent' speakers of a language. I shall assume that in order to understand something in this sense, a speaker must be able on reflection to articulate, or recognize, without further instruction correct explications of the meaning, at least in those cases of meaning where explications are relevant to understanding. So, for example, those who understand the conventional linguistic meaning of 'chair' could on reflection articulate or at least recognize as correct a correct dictionary explication.

I should add that Frege's conception of sense is distinct not only from conventional linguistic meaning, but also from modern conceptions of idiolectic linguistic meaning. I shall concentrate on conventional linguistic meaning, because (except in the case of proper names and indexicals) Frege's notion of sense is much more likely to be confused with a conception of meaning that has social elements. But my discussion of linguistic meaning is meant to be neutral, except where the context indicates otherwise, on issues of the relative roles of individual and community in fixing meaning. Frege's notion of sense contrasts with all modern conceptions of linguistic meaning, including those concerned with meaning in idiolects. Of course, since he conceived of sense as a sort of linguistic 'meaning', sense is something with aboutness properties which is expressed through language. To this degree, Frege's conception of sense is a conception of linguistic meaning. So when I write of 'linguistic meaning' with the intent of contrasting such a notion with Frege's notion of sense, I am using this notion in a way that associates it with modern conceptions.

II

As background for what ensues, one must bear in mind Frege's repeated claim that his primary concern in producing his logical theory is to theorize about thought and truth, not language. He says that language often obscures the actual structure and nature of thought. Of course, he also held that one could not think most of the thoughts one thinks except by means of language. So, as he says, he is forced to concern himself with language even though it is not his main interest. Thus natural language is cast in the role of a villain on which one must perforce rely, both in order to think and as a source of clues about the nature of thought. But natural language is an imperfect instrument for thought. And Frege believed that only a language yet to be fully fashioned—a 'perfect language' (one ideally suited to the expression of thought, especially *a priori* thought)—would express thought, and senses, in a perspicuous manner.

One might be inclined to assimilate this point of view to more modern conceptions according to which the form and semantical characteristics of a language can be understood only by reference to its 'deeper structure'. There is something to be said for this interpretation. But taken in the way it would normally be understood, it underestimates how far removed, for Frege, thought and sense may be from a language's surface.

I shall be arguing that no amount of investigation of the actual usage and understanding of one's language is, according to Frege, sure to reveal the 'deeper structure' and nature of sense and thought. One may have to achieve genuine advances in non-linguistic knowledge before one can fully master the senses and thoughts that one's language expresses. This point renders Frege's view and modern methodologies for studying language quite different.[5]

[5] Another background point that serves as a clue to Frege's distinction between sense and conventional linguistic meaning lies in his remarks about proper names, demonstratives and other indexicals. He ascribes sense to names, but denies that there need be any conventionally agreed-upon sense. He frequently says that the sense of expressions like 'I', 'now' and 'yesterday' shift with the context and referent. Clearly their linguistic meaning, the accepted norm for conventional and idiolectic use and understanding of the language, remains the same through these shifts. Conversely, Frege sometimes counts the sense associated with two applications of two different indexical expressions ('yesterday' and 'today') the same, whereas the linguistic meaning is clearly different. These remarks are obviously incompatible with taking the sense of these expressions to be their conventional linguistic meanings.

It would be easy, though slovenly, to think of proper names and demonstrative expressions as special cases and to dismiss Frege's remarks about language getting in the way of thought as typical expressions of a logician interested in regimentation. But there is much more behind these remarks. They are symptoms of a radically epistemic conception of sense.

These background points are discussed at some length in 'Sinning Against Frege' (Ch. 5 above), in my Review of Dummett's *The Interpretation of Frege's Philosophy, The Philosophical Review*, 93 (1984), 454–8, and in 'Frege on Extensions of Concepts' (Ch. 7 below). The latter paper contains other grounds for distinguishing conventional linguistic meaning and sense. The present paper is a development of points first made in 'Frege on Extensions of Concepts'.

We can begin to appreciate the full extent of the difference between sense and conventional linguistic meaning by developing Frege's statements about understanding and about *vagueness*.

We enter this subject by posing an interpretative puzzle. The puzzle consists of an argument that leads from Frege's statements and a plausible further premise to an inconsistency. Resolving the inconsistency as Frege would will, I think, deepen our understanding.

The puzzle arises out of the conjunction of three claims that Frege makes. The first is that vague expressions lack a *Bedeutung* (e.g. *PW* 112, 155, 179/ *NS* 133, 'Ausführungen über Sinn und Bedeutung' (1892–1895); *NS* 168, 'Begründung meiner strengeren Grundsätze des Definierens' (1897/8 oder kurz danach); *NS* 193–4, 'Über Schoenflies: die logischen Paradoxien der Mengenlehre' (1906); *PMC* 114/ *WB* 183, Frege to Peano, 9/29/1896; 'About the Law of Inertia', *Synthese*, 13 (1961), 360–1/ *KS* 122–3; O 158–160). Frege states this principle repeatedly. We need not discuss his reasons.

The second is that all of the expressions in conventional mathematics and natural science, and some of the same terms in ordinary discourse, are not sharply understood—are not backed by a sharp grasp of a definite sense—even by their most competent users (*PW* 221–2, 216–7, 211/ *NS* 239–40, 234, 228, 'Logik in der Mathematik' (1914); *FA* vii in conjunction with *FA* sections 1, 2; *FA* 81; cf. also 'About the Law of Inertia', trans. Rand, *Synthese*, 13 (1961), 360/ *KS* 122; O 158).

If the best mathematicians lack such a grasp, one might reason, conventional mathematical usage is surely, by Frege's lights, vague. And if mathematical expressions are thus vague, surely most other expressions are too. This is the premise I foreshadowed.

The third element in the puzzle is Frege's repeated citations of this or that expression from mathematics or ordinary discourse as having such and such a *Bedeutung*. Indeed, Frege consistently writes as if lacking a *Bedeutung* is an aberration, except in fiction.

These points must be properly appreciated in order to assess the relevance of attacks on Frege over the last two decades. It has been common to hold that Frege had mistaken views about the 'semantics' of proper names, demonstratives, and the like. Sometimes, with a dutiful nod toward 'the historians', a critic will say that 'Fregeans though perhaps not Frege' held certain mistaken views. I think that the historical issue is of more than historical importance. Frege held an entirely different conception of 'semantics' than do most of his modern critics. His conception of sense was introduced to serve this conception. Language was, in his view, an imperfect vehicle for thought; and thought itself was understood in a highly idealized manner that I shall try to articulate. By contrast, modern conceptions of language concentrate on explaining linguistic competence—what is standardly or normally understood by linguistic expressions.

If the philosophical interest and importance of modern conceptions of the 'semantics' of referring expressions is to be reasonably evaluated, one must be clear about what one is doing when one gives a 'semantical account' of the language. By paying too little attention to the range of possible conceptions of language (or semantics), modern critics of Frege often not only underestimate their target but fail to appreciate the nature of the philosophical problems that motivated his theory of sense. Since these problems remain philosophically important, discussions of them that fail to appreciate how they are understood in Frege's own work almost inevitably suffer substantively, as well as historically.

The three claims, together with the reasoning following the second, are jointly inconsistent. For if one holds that even the best mathematicians do not have a firm grasp of a sense that fixes a sharply bounded *Bedeutung* for concept words and concludes from this that most mathematical and other (concept) expressions are vague; and if one holds that vague concept expressions lack *Bedeutung*; one cannot go about assigning *Bedeutungen* to most of the concept expressions of mathematics and ordinary discourse. Here we have a puzzle. Did Frege find his way through it?

III

Although the puzzle centers on *Bedeutung*, much of its interest bears on sense. So in discussing it, we shall keep track of both notions. In arriving at an interpretation, one must bear in mind the fact (not seriously questionable, I believe) that Frege was committed to the view that senses of concept expressions fix definite, sharp *Bedeutungen* where they have any *Bedeutung* at all. Frege conceived of senses as eternal abstract entities that have their logical and semantical properties independently of the activity of actual minds or language-users. Neither senses themselves nor their applications to *Bedeutungen* could be vague. This in itself constitutes a difference between Frege's conception and the ordinary conception of conventional linguistic meaning. But this difference will stand out in sharper relief once we have unravelled the puzzle.

Let us begin by considering whether one could avoid attribution of the third of the three claims. One possible line would be this: 'Despite what Frege says in illustrating his doctrines, he does not really mean that our expressions have definite senses or *Bedeutungen*. For if they are vague, they do not, by his own account, have *Bedeutungen*; and if the most competent users of the expressions do not fix definite *Bedeutungen* through their linguistic practices, the expressions would seem also to lack definite senses. The fact that Frege does not speak this way could be attributed to his desire not to complicate his expositional task unduly. He conveyed his idealized picture of sense and *Bedeutung* through his examples. But he did not see imperfect natural language as embodying his idealization.'

There are two options within this general line. We might say that Frege thought that ordinary expressions in mathematics and elsewhere had a sense but lacked a *Bedeutung*. Or we might say that he thought such expressions lacked a sense as well. Let us take up the second option first.

There is an early passage in *Begriffsschrift* (1879) (*B* 27) where Frege says of a particular vague expression that it lacks a 'judgeable content'. But he does not indicate how far he intends his remark to be generalized. And since 'content' is a technical expression that Frege later thought blurred the sense–*Bedeutung* distinction, one cannot rest much weight on the remark. More significantly, there is a letter to Peano (1896) that reads as follows:

> The fallacy known by the name of 'Acervus' rests on this, that words like 'heap' are treated as if they designated a sharply bounded concept whereas this is not the case. Just as it would be impossible for geometry to set up precise laws if it tried to recognize threads as lines and knots in threads as points, so logic must demand sharp limits of what it will recognize as a concept unless it wants to renounce all precision and certainty. Thus a sign for a concept whose content does not satisfy this requirement is to be regarded as without *Bedeutung* from the logical point of view. It can be objected that such words are used thousands of times in the language of life. Yes; but our vernacular languages are also not made for conducting proofs. And it is precisely the defects that spring from this that have been my main reason for setting up a conceptual notation. The task of our vernacular languages is essentially fulfilled if people engaged in communication with one another connect the same thought, or approximately the same thought, with the same proposition. For this it is not at all necessary that the individual words should have a sense and *Bedeutung* of their own, provided only that the whole proposition has a sense. Where inferences are to be drawn the case is different; for this it is essential that the same expression should occur in two propositions and should have exactly the same *Bedeutung* in both cases. It must therefore have a *Bedeutung* of its own, independent of the other parts of the sentence. In the case of incompletely defined concept words there is not such independence: what matters in such a case is whether the case at hand is the one to which the definition refers, and that depends on the other parts of the proposition. Such words cannot therefore be acknowledged to have an independent *Bedeutung* at all. This is why I reject conditional definitions of signs for concepts. (*PMC* 114–115/ *WB* 183, Frege to Peano, 9/29/1896)

This passage raises a number of interpretative problems that I will not go into. It could perhaps be read as supporting the view that Frege thought that virtually all words in ordinary discourse lack a definite sense. Perhaps each would be attached to a cluster of senses, no one of which it definitely expresses. But the passage certainly does not demand this interpretation. There is nothing in the passage to indicate that Frege is assimilating virtually *all* words in natural discourse to 'heap'.

The main point of the passage is that one must insist on explicit rather than contextual definitions. Frege is laying down a requirement about how we should analyse language for the purposes of logic (as contrasted with the purposes of the vernacular). Similar requirements apply to the project of defining 'number' (cf. *FA*, middle sections). But Frege repeatedly assumes that 'number' has a definite sense and denotation. For example, he writes of a certain large finite numerical expression that it has a 'perfectly definite sense' (*FA* 114).[6] Words like 'heap' seem to be treated as special cases, not as paradigms of all linguistic usage.

[6] This occurrence of 'sense' (*Sinn*) is not corrected in the letter to Husserl of May 24, 1891, that explicitly corrects a number of passages in *Foundations of Arithmetic* that blur the sense–*Bedeutung* distinction, cf. *PMC* 63/*WB* 96.

It should be noted that this passage suggests a distinction between the notion of sense and the notion of conventional or idiolectic linguistic meaning. The implication of the passage is that the relevant vague notions, like 'heap', do not have a definite sense or denotation. Different people in different contexts, or the same person in different contexts, may associate different thoughts with sentences containing the word, in such a way that there is no saying what thought component is the sense associated with the word. Sometimes such thought components will apply to 'the case at hand' (say, counting it a heap); other times, they will not. One cannot associate any particular sense with the relevant vague word on all its occasions of use, though it may have different senses on different occasions.

There is no suggestion that such words are strictly without what we would call meaning. They contribute to the fulfilment of the 'task of our vernacular languages' by facilitating communication. It might well be that despite variations in putative extension among language-users (or by a given user at different times), one could attribute a constant meaning—a constant norm for understanding—to the word. I see no reason to think that Frege would have denied that an attribution of vague linguistic meaning was possible, based on an account of vague linguistic usage. Frege's own analogy suggests that he would not have denied this. There is no suggestion that one couldn't give a theory of knots and threads. It just would not be directly relevant to geometry. Similarly, an account of vague usage would not be an account of sense.

Let us return to the option of solving the puzzle by denying that most ordinary and mathematical expressions have a definite sense. The primary trouble with the option is that there is no evidence that Frege thought that nearly all the expressions we actually use lack a definite sense. In fact, he often seems to suggest that it is relatively easy to get an expression to express a sense. He writes:

> If one is concerned with truth . . . one has to throw aside concept words that do not have a *Bedeutung*. These are . . . such as have vague boundaries . . . [But] the context cited need not lack a sense, any more than other contexts in which the name 'Nausicaa', which probably does not denote or name anything, occurs. But it behaves as if it names a girl, and it is thus assured of a sense. (*PW* 122/ *NS* 133, 'Ausführungen über Sinn und Bedeutung' (1892–1895))

Here it is clear both that senses are easily attached to words and that vague concept words do not, or need not, lack sense. It is also at least suggested that the vague words comprise a relatively small, delimited class.[7]

Frege argues in numerous places that the objectivity of science and the possibility of communication depend on sentences' expressing (objective) sense (e.g. G & B 46/ *KS* 170; 'Über Begriff und Gegenstand', O 196n; *PMC*

[7] Two pages later Frege writes that not being 'an empty sequence of sounds' is sufficient for a proper name's having a sense.

80/ *WB* 128, Frege to Jourdain, undated, *c*.1913; *LI* 32 ff., 24 ff./ *KS* 362 ff., 'Die Verneinung', O 143 ff.; *KS* 358 ff.; 'Der Gedanke'; O 74 ff.). And he repeatedly states or implies that the sense of a sentence (at an occurrence) is a function of the senses of its parts (G & B 54/ *KS* 178, 'Über Begriff und Gegenstand', O 205; *PW* 192/ *NS* 208–209, 'Einleitung in die Logik' (1906); *LI* 55 ff./ *KS* 378 ff., 'Gedankengefüge', O 37 ff.).

These general doctrines about the commonness of sense-expression are buttressed by numerous statements that this or that expression has a sense: 'the morning star', 'the celestial body most distant from the Earth', 'the least rapidly convergent series', 'the Moon', 'Odysseus' and so on are cited in 'On Sense and Denotation' alone. In fact, the general tenor of Frege's remarks is that only rationally defective expressions (like 'the concept horse', as he later decided (*PW* 177–8/ *NS* 192–3, 'Über Schoenflies: die logischen Paradoxien der Mengenlehre' (1906); *PW* 193/ *NS* 210, 'Einleitung in die Logik' (1906)), really lack a sense. Although the examples I have cited concern singular terms, many of the cited singular terms contain concept words. So by the composition doctrine, the singular terms could not have senses unless their component concept words had senses. Moreover, there are passages in which Frege explicitly speaks of concept words in conventional usage as having a definite sense (e.g. *PW* 122, 209/ *NS* 133, 'Ausführungen über Sinn und Bedeutung' (1892–1895); *NS* 225–6, 'Logik in der Mathematik' (1914)). The view that nearly all expressions lack a definite sense has no clear textual basis and runs against an overwhelming number of remarks Frege makes.

We turn now to the option of taking Frege to hold that nearly all expressions in ordinary, conventional mathematics and ordinary discourse lack a denotation without lacking a sense. This option might seem to receive support from the following unpublished passage written in 1914:

> In the first stages of any discipline we cannot avoid using the words of our language. But these words are, for the most part, not really appropriate for scientific purposes, because they are not precise enough and fluctuate in their use. Science needs technical terms that have precise and fixed *Bedeutungen*, and in order to come to an understanding about these *Bedeutungen* and exclude possible misunderstandings, we give explications. Of course, in so doing we have again to use ordinary words, and these may display defects similar to those which the explications are intended to remove. So it seems that we shall then have to do the same things over again, providing new explications. Theoretically one will never really achieve one's goal in this way. In practice, however, we do manage to come to an understanding about the *Bedeutungen* of words. Of course we have to be able to count on a meeting of minds, on others guessing what we have in mind. (*PW* 207/ *NS* 224, 'Logik in der Mathematik' (1914))

The practice of substituting technical expressions for ordinary expressions, even those like 'number', is prominent in *Grundgesetze*, written much earlier.

Here Frege does perhaps suggest that nearly all non-technical words are 'not precise enough and fluctuate in their use'.

Nevertheless, Frege does not say in this passage that non-technical words, much less ordinary mathematical expressions, are in general bereft of *Bedeutung*. In fact, the passage clearly presupposes that the relevant words *have a Bedeutung*. The problem that concerns him is that there is some difficulty in determining or understanding exactly what *Bedeutung* a person denotes with an ordinary word. The problem is lack of standardization. The remarks of this passage appear to apply primarily to expressions of ordinary language. And Frege is issuing a general indictment, not one specifically concerned with vagueness. Imprecision and fluctuation of use cover not only vagueness but ambiguity; lack of interpersonal standardization; lack of a generally agreed-upon specification, in other terms, of the *Bedeutung*; and the dependence of expressions (demonstratives or indexicals) on context.

The problem with this option for solving the puzzle is similar to the problem with the previous option. Frege repeatedly writes as if fiction and serious mistakes are the primary sources of failures of denotation. When he discusses vague, denotationless concept words, the examples he offers are always presented as if they were special cases—'heap', 'bald' and so on. By contrast, Frege gives numerous examples of concept words—both in conventional mathemataies and in ordinary discourse—that are explicitly ascribed denotations (*PW* 120–121, 177, 182, 229/ *NS* 130–131, 'Ausführungen über Sinn und Bedeutung' (1892–1895); *NS* 192, 198, 'Über Schoenflies: die logischen Paradoxien der Mengenlehre' (1906); 247, 'Logik in der Mathematik' (1914); G & B 113/ *KS* 278, 'Was ist eine Funktion?', O 663; and so on).

If he were to hold that nearly all concept words in mathematics, natural science, and ordinary discourse are vague, and thus lack a *Bedeutung*, he would have to hold that nearly all sentences in actual use strictly speaking express neither truths nor falsehoods. There is no suggestion of a 'secret doctrine' to this effect. (The closest approach to it that I can find is *PW* 242–243/ *NS* 261–262, 'Logik in der Mathematik' (1914)). Indeed, again, his basic argument for the existence of sense depends on the assumption that we share knowledge, hence express true thoughts, in the sciences. Neither of the options that we have considered for denying the third step in the argument is a credible expression of the spirit or letter of Frege's writings.

There is no questioning the attribution of the first step of our argument—the claim that vague expressions lack a *Bedeutung*. Frege's remarks are frequent, straightforward and explicit. What of the second step—the claim that expressions of mathematics, science and ordinary discourse are not sharply understood, even by their most competent users? Understanding it is a complex undertaking—the central element in solving our puzzle. I shall be developing an interpretation of the second step throughout the rest of the paper.

IV

On several occasions, Frege states or implies that the sense of an expression, often a mathematical expression, is not clearly or sharply grasped even by the most competent users of the expression. This theme emerges most clearly in his late writing. But the idea informs Frege's work almost from the beginning. Let us start by considering some passages from *Foundations*:

What is known as the history of concepts is a history either of our knowledge of concepts or of the meanings of words. Often it is only through great intellectual labour, which can continue over centuries, that a concept is known in its purity, and stripped of foreign covering that hid it from the eye of the intellect. (*FA* p. vii)

A cognate passage occurs seven years later, in 1891:

For the logical concept, there is no development, no history . . . If instead of [this sort of talk] one said 'history of the attempt to grasp a concept' or 'history of the grasp of a concept', it would seem to me far more appropriate; for the concept is something objective that we do not form and is not formed in us, but that we try to grasp and finally, it is hoped, really grasp—if we have not mistakenly sought something where there is nothing. (*KS* 122, 'Über das Trägheitsgesetz', O 158)[8]

The significance of these two passages can be appreciated only when the context of the first one is noted. That passage occurs in the introduction to *The Foundations of Arithmetic*. It is almost immediately followed in sections 1 and 2 by a long complaint that most of the fundamental notions of arithmetic have not been 'sharply determined' or 'sharply grasped' (section 1) by mathematicians, past or present. He cites the notions of function, continuity, limit, infinity, negative and irrational numbers.

The whole book is thus cast as an attempt to produce, for the first time, a sharp grasp of concepts (later, senses and concepts through senses) that are associated with conventional mathematical usage, but that had never before been 'sharply grasped'. In the course of the book, Frege gives several further examples of applications of mathematical expressions that past usage and understanding had left indeterminate, but which his own understanding purports to fix determinately (*FA* 68, 80–81, 87, 96–98, 100–101, 102–103).

The mist metaphor recurs thrice in 1914. Here is one occurrence:

How is it possible, one may ask, that it should be doubtful whether a simple sign [in common use] has the same sense as a complex expression if we know not only the sense of the simple sign, but can recognize the sense of the complex one [that

[8] The first of these passages was written before the development of the sense–*Bedeutung* distinction (1891). It is likely but not certain that the second one was also. The talk of 'grasping' suggests, however, that the passages concern what is thought. Frege writes in 1897: 'We might cite, as an instance of thoughts being subject to change, the fact that they are not always immediately clear. But what is called the clarity of a thought in our sense of this word is really a matter of how thoroughly it has been assimilated or grasped, and is not a property of a thought' (*PW* 138/*NS* 150, 'Logik' (1897)).

purportedly analyses it] from the way it is put together? The fact is that if we really do have a clear grasp of the sense of the simple sign, then it cannot be doubtful whether it agrees with the sense of the complex expression. If this is open to question although we can clearly recognize the sense of the complex expression from the way it is put together, then the reason must lie in the fact that we do not have a clear grasp of the sense of the simple sign, but that its outlines are confused as if we saw it through a mist. The effect of the logical analysis of which we spoke will then be precisely this—to articulate the sense clearly. (*PW* 211/ *NS* 228, 'Logik in der Mathematik' (1914))

A second occurrence appears in a critical discussion of ordinary mathematical practice. Frege makes the point that ordinary mathematicians often grasp the same senses. In particular, they 'attach the same sense to the word "number" '. But because they are so cavalier about their definitions, they do not 'get hold of the sense properly'. They do not manage to attain a definite mastery of the sense they think with. He says that the sense appears to them

in such a foggily blurred manner that when they make to get hold of it, they reach for it in the wrong place. One reaches perhaps erroneously to the right, the other to the left; and so they do not get hold of the same thing, although they wanted to. How thick the fog must be for this to be possible! (*PW* 216–217/ *NS* 234, 'Logik in der Mathematik' (1914); the third occurrence of the mist metaphor occurs later in the same essay, *PW* 241–242/ *NS* 260–261, and will be mentioned below).

The comedy of this passage is double-edged. Frege produces a deliciously absurd picture of his opponents. But the terms of the metaphor indicate a notion of understanding on Frege's part (as grabbing some elusive phantom) that inevitably seems incongruous to modern minds.

Frege goes on to blame the failure of mathematicians to 'grasp' or 'get a proper hold' of the senses that they attach to the word 'number' on a failure to lay down and abide by good definitions. There follows a lengthy discussion of Weierstrass's failures to provide clear definitions for his primary mathematical expressions. Frege is obviously taking Weierstrass as an example of one of the mathematicians who 'attach the same sense to the word number', but do not get hold of it properly. Frege thinks Weierstrass does not grasp sharply the sense that he associates with his own word. Frege ends this discussion with the remark that Weierstrass's grasp of the notion of number is very unclear. And he diagnoses the trouble as follows: 'He lacks the ideal of the system of mathematics' (*PW* 221/*NS* 239 'Logik in der Mathematik' (1914)).

The obvious implication of these passages is that the most competent users of the relevant expressions may not have achieved a sharply bounded understanding of a sense—a 'sharp grasp' of a definite sense—that they attach to the word they use. Conventional usage and understanding may not give those users a clear understanding of a definite sense that the expression already expresses.

From this remark it is natural for us to infer that the relevant expressions are vague. For if conventional use and understanding *by even the most* *competent users* does not constitute mastery of a sense that fixes sharp boundaries, it is natural to conclude that the expression's 'sense' and proper application simply lack sharp boundaries. Attributing this inference to Frege is, however, the mistake that leads to our puzzle.

For Frege did not draw this inference. The clear implication of the above cited passages is that the relevant expressions *have* a definite sense, or denote a definite concept. The problem is that no-one has yet grasped it fully or clearly. The defect is in our understanding and use, not in the abstract senses themselves, nor even in our expressions' relations to their senses. There is, moreover, no suggestion that such senses lack a sharply delimited denotation.

Thus, although Frege believed something that for most modern philosophers would appear to entail that the senses of most mathematical expressions are vague, he did not accept the entailment. He might well concede that what we would call an expression's 'conventional linguistic meaning', as fixed by conventional or actual usage, is vague. (Recall the knots and threads passage.) But he thought that the expression might nevertheless have a definite *sense* with a sharply bounded denotation.

Now this fact, not the mere resolution of the puzzle, is the primary point of interest. Frege's views about vagueness are in themselves not particularly important. They are significant almost entirely because they are symptomatic of interesting aspects of his conception of thought and sense: Frege's failure to draw the inference is, I think, a sign that his conception of sense differs substantially from any ordinary conception of conventional linguistic meaning.

For all the controversy that has surrounded the term 'linguistic meaning' as a theoretical tool, I think it clear that we have a rough and ready, reasonably coherent conception that is expressed in standard dictionary entries and in intuitively acceptable explications of the meanings of terms. Conventional linguistic meaning, according to this ordinary conception, is a complex idealization of conventional use and understanding. The meaning of a term is revealed in its use and is articulated in dictionary entries and in reflective explanations of its use by competent users. If conventional use and best understanding are not sharp, then conventional meaning is certainly not sharp. An analogous point can be made for the notion of idiolectic meaning. On this common conception, it would be absurd to say that an expression had a definite, non-vague conventional linguistic meaning—or, alternatively, a definite range of application—even though all extant understanding and all actual usage failed to fix sharp boundaries for the expression's application. Similar points hold for modern conceptions of idiolectic linguistic meaning. It would be absurd to say that an expression had a definite, non-vague idiolectic linguistic meaning, even though all the individual's abilities to articulate the expression's meaning, and all the individual's actual usage failed to fix sharp boundaries for the expression's application. (Cf. n. 4.)

So it appears that for Frege the capacity of a word to express a sense is partly independent of the user's understanding and use of the word, and even of conventions about the word's usage, in a way that distinguishes his notion of sense from almost any ordinary notion of linguistic meaning—conventional or idiolectic.[9]

V

Before proceeding, I want to note that Frege does not explicitly emphasize the contrast with conventional (or idiolectic) linguistic meaning that I am highlighting. I believe that he never thought the issue through in any depth. The view that I have attributed represents a prominent strand in his thinking. And I think that he never gave it up. But there are passages that suggest that Frege was not focusing on the matter, passages in which some further specification of the view would be natural had Frege clearly formulated it to himself. There are also passages which suggest that Frege wanted to avoid questions that this conception very naturally raises.

According to the conception that we have been developing, the sense and *Bedeutung* of an expression in conventional mathematics might be quite definite even though all past and present applications and all extant abilities to explicate a word fail to fix sharp boundaries. Frege indicates, however, that the *Bedeutung* of a mathematical expression sometimes changes in the course of its history. Frege indicates in 'Function and Concept' that the *Bedeutung* of the word 'function' has changed in the course of the history of mathematics. More is now included in the *Bedeutung* than had been earlier, Frege counts inclusion of transition to the limit as an essential broadening of the original *Bedeutung* of 'function'. And he counts the admission of new arguments and values (such as negative numbers) for standard mathematical operations as evidence that the *Bedeutung* of the expression has been extended.

A similar point applies to the expression 'number'. In 1914 he writes:

> Originally the numbers recognized were the positive integers, then fractions were added, then negative numbers, irrational numbers, and complex numbers. So in the course of time wider and wider concepts came to be associated with the word 'number'... And the same happened with other arithmetical signs. This is a process which logic must condemn and which is all the more dangerous, the less one is aware

[9] It is important to recognize the difference between this distinctive feature, which bears on the relation between senses and words, and two other striking features of Frege's doctrine of sense. Frege conceived of senses as eternal abstract entities that have their logical and semantical properties independently of the activity of actual minds. And he held that the content of all thinking could be characterized purely in terms of such entities. These latter two views, though perhaps congenial with the conception that we are discussing, are strictly independent of it. Of the three views, I find the first (the one I am concentrating on in this paper) promising, the second deeper and more interesting than commonly thought but extremely dubious, and the third mistaken. For criticism of the third view, see 'Belief *De Re*', and 'Sinning Against Frege' (Ch. 5 above).

of the shift taking place. The history of science runs counter to the demands of logic . . . No science can master its subject matter and work it up to such transparency as mathematics can; but perhaps also, no science can lose itself in such thick mist as mathematics if it dispenses with the construction of a system. As a science develops, a certain system may prove no longer to be adequate, not because parts of it are recognized to be false but because we wish, quite rightly, to assemble a large mass of detail under a more comprehensive point of view in order to obtain greater command of the material and a simpler way of formulating things. In such a case we shall be led to introduce more comprehensive, i.e. superordinate, concepts and relations. What now suggests itself is that we should, as people say, extend our concepts. Of course . . . we do not alter a concept; what we do rather is to associate a different concept with a concept word . . . The sense does not alter, nor does the sign, but the correlation between sign and sense is different . . . If we have a system with definitions that are of some use and aren't merely there as ornaments, but are taken seriously, this puts a stop to such shifts taking place . . . In fact, we have at present no system in arithmetic. All we have are movements in that direction. Definitions are set up, but it doesn't so much as enter the author's head to take them seriously and to hold himself bound by them. So there is nothing to place any check on our associating, quite unwittingly, a different *Bedeutung* with a sign or word. (*PW* 241–242/*NS* 260–261, 'Logik in der Mathematik' (1914))

This passage is not incompatible with the passages I previously centered upon. Indeed they are complementary. This last one occurs in the same essay as the most explicit remarks about the best mathematicians' lacking a clear grasp of the senses they all attach to terms they commonly deal with. The mist metaphor even recurs in this passage.

But seen very schematically, the two groups of passages suggest different models of progress in science. On the first model, progress is a matter of obtaining a better, clearer grasp of thoughts that one is already dimly thinking and unperspicuously expressing. Better theory results in deeper understanding and clearer explication of some of one's own thoughts and senses. One might have been thinking various thoughts at different times, even though one is using a single expression. But deeper theory cuts through the mist and explicates a given sense long associated (however inconstantly and uncomprehendingly) with an expression. On the second model, theoretical improvements issue in new connections between words, on one hand, and senses and *Bedeutungen*, on the other. One replaces old concepts (and senses determining them) with new ones in one's thinking.

If the first model is to have any place at all, it must be complemented with the second. Sometimes actual usage is just too far removed from some conceptions to attribute those conceptions to the users. When the users eventually arrive at such a conception, and express it in old words, we have to see them as having altered the senses of the words. On the other hand, this point is compatible with the idea that *sometimes* a user expresses a conception that is more definite and richer than the users could articulate given his or her state of knowledge.

There remain vexed questions of practice and principle about where to draw the line. I believe that Frege never concentrated on the differences between the two models, or worried about the issues regarding change-of-belief *versus* change-of-'meaning' that have become prominent since his time. As noted, some of Frege's discussion seems fashioned so as to lay these issues aside.[10] I shall, however, continue to concentrate on the first model on the ground that it represents something philosophically important and distinctive, even if undeveloped, in Frege's thinking.

VI

How *could* an expression express a definite sense if its being related to that sense were not entirely explicable in terms of how its users actually use and understand the expression?[11] Frege does not directly confront the question. But his work seems to indicate a certain type of answer. Saying that the expression 'Number' denotes a definite concept and expresses a definite sense, despite the insufficiencies of current understanding and usage, is (for Frege) made possible by the belief that the ultimate justification of current mathematical practice supplements current usage and understanding in such a way as to explicate a definite concept and a definite sense that no one may currently be able to thoroughly understand or articulate. Weierstrass's inadequate understanding of the senses he thought with would be rectified if the 'ideal of the system of mathematics' had properly informed his thinking.

Frege's conception of sense-expression has two fundamental presuppositions. The *first* is that mathematics and other cognitive practices are founded on deeper rationally understandable aspects of reality than anyone may have

[10] For example, Frege holds that for systematic purposes it is better to introduce new mathematical signs for old ones when one gives a new definition. He largely follows this practice in *Basic Laws*. At least once he writes of a given concept expression that 'as long as it remains incompletely defined [definiert]' in a way that fails to determine for every object whether that object falls under the denoted concept, it 'must remain undecided' whether the object falls under the concept. So the concept expression is vague (*PW* 242–3/ *NS* 261–2, 'Logik in der Mathematik' (1914)). Perhaps Frege means only that as far as the incomplete definition goes, the word is left vague. Literally read, however, this passage is incompatible with the various passages, cited earlier, in which Frege speaks of the denotations of various mathematical concept expressions in conventional mathematics, even though no-one had given those expressions 'complete' definition. Frege is clearly more interested in giving an ideal system of mathematics than in providing a systematic account of the state of conventional mathematics before the ideal system has been discovered.

[11] This sort of question is something that Wittgenstein pressed in his thinking about rule-following. It is also implicit in Dummett's insistence that a theory of 'meaning' be a theory of understanding and use. I think that the question is legitimate and profoundly difficult, though I think that Dummett is not very sensitive to the elements in Frege that ignore this insistence. Indeed, he often appears to attribute acquiescence in the insistence to Frege.

presently understood. The implications of a practice can reach beyond the procedures and dispositions of the practitioners.[12]

The *second* presupposition is that mastery of these deeper rationales is to be treated as involving insight into the true senses of expressions. It is to be understood in terms of conceptual clarification. Deeper insight into the nature of things is simultaneously deeper insight into rational modes of thinking about the nature of things.

Both presuppositions derive from the rationalist tradition. They combine to form a rather special conception not only of the nature of sense-expression, but also of the enterprise of philosophical analysis. I mean this latter phrase not in a narrow sense that would apply specially to 'analytic' philosophy—the tradition that grew out of Frege and came in our century to dominate the English-speaking philosophical world. I mean by 'philosophical analysis' something that would apply to much of the philosophical activity of Socrates, Plato, Aristotle, Descartes, Leibniz and Kant, as well as to twentieth-century 'analytic' philosophers more specially concerned with language. Any activity at least partly aimed at understanding our 'conceptual scheme', our cognitive practices, is analytic in this broad sense. Frege shared with the rationalist tradition the confidence that a deep rationale underlies many of our practices. He also shared the view that this rationale makes its imprint on those practices in such a way that understanding the rationale is achieving insight into what was part of the practices even before the understanding was achieved. Since Frege develops the second of the two cited presuppositions in some depth, at least implicitly, I shall begin by discussing it. I shall return to the first afterward.

There are in Frege's writings two contrasting themes about the grasping of senses. Frege does not explicitly juxtapose and reconcile them until relatively late in his career (1914). But both themes and the makings for reconciliation are present from early on. One theme emphasizes how relatively easy it is to express and grasp a sense. We have already noted a variety of passages in which Frege suggests that normal linguistic activity is often sufficient to *express* a sense. Frege also states (G&B 57–58/ *KS* 144; 'Über Sinn und Bedeutung', O 27) that a sense expressed is grasped by everyone 'sufficiently familiar with the language'. He does not say what constitutes sufficient familiarity. But the surrounding context certainly suggests that the requisite familiarity is not an exceptional accomplishment. In the same passage, Frege cites names and (presumably) other context-dependent devices as exceptions

[12] Interpreting this point requires delicacy. Frege was primarily interested in mathematics. And he was pursuing a foundationalist programme. I think that his conception of sense should not be tied too closely to these contingencies. Frege clearly intended to apply his conception of sense to domains other than the mathematical. And neither foundationalism nor reductionism are crucial to the conception. Sense-expression is to be explicated in terms of better rationales than anyone may have grasped—regardless of whether these rationales provide a foundation in anything like the way logicism was supposed to provide a foundation for mathematics.

to the remark that knowing the language is sufficient to grasp the sense. But he indicates that people nevertheless commonly 'attach' a sense to the relevant term in a context, presumably grasping it in so doing.

There are also numerous passages in which Frege says that thinking paradigmatically involves grasping a thought, grasping the sense of a declarative sentence. There is no suggestion that thinking is especially difficult to engage in. (Frege's own virtuosity in this regard may have blinded him to difficulties that the rest of us face!) In fact, Frege bases one of his arguments for the existence of senses on the assumption that mankind commonly grasps them in common. All of these passages indicate that senses are routinely grasped in the hurly-burly of cognitive life.

On the other hand, there are passages where Frege indicates that grasping a sense or thought is a matter of degree, and that thoroughly grasping thoughts is an achievement worthy of some renown. There are, for example, the two passages from 1884 and 1891 in which Frege says that what is usually called the history of a concept is better termed the history of attempts to grasp a concept (*FA* p. vii; *KS* 122, 'Über das Trägheitsgesetz', O 158). Frege elaborates the point in a passage from 1897 (*PW* 138/ *NS* 150, 'Logik' (1897)). He is arguing that thoughts are changeless, in precise analogy to the earlier arguments that 'concepts' do not change: 'We might cite, as an instance of thoughts being subject to change, the fact that they are not always immediately clear. But what is called the clarity of a thought in our sense of this word is really a matter of how thoroughly it has been assimilated or grasped, and is not a property of a thought.' Taken together, these remarks suggest that 'thorough grasp' is a difficult matter.

The suggestion is amplified in Frege's late writing (1914). Where there is a failure to provide a satisfactory explication of mathematical terms, there is, he writes, a failure to understand the terms. Writing of Weierstrass, Frege says in an unusually smug passage:

> He had a notion of what number is, but a very unclear one; and working from this he kept on revising and adding to what should really have been inferred from his definition... And so he quite fails to see that what he asserted does not flow from his definition, but from his inkling of what number is. (*PW* 221/*NS* 239, 'Logik in der Mathematik' (1914))

He continues:

> But how, it may be asked, can a man do successful work in a science when he is completely unclear about one of its basic concepts? The concept of a positive integer is indeed fundamental for the whole arithmetical part of mathematics. And any unclarity about this must spread throughout the whole of arithmetic. This is obviously a serious defect and one would imagine that it could prevent a man from doing any successful work whatsoever in this science. Can any arithmetical sentence have a completely clear sense to someone who is in the dark about what a number is? This question is not an arithmetical one, nor a logical one, but a psychological one. We

simply do not have the mental capacity to hold before our minds a very complex logical structure so that it is equally clear to us in every detail. For instance, what man, when he uses the word 'integral' in a proof, ever has clearly before him everything which appertains to the sense of this word! And yet we can still draw correct inferences, even though in doing so there is always a part of the sense in penumbra. Weierstrass has a sound notion of what number is and working from this he constantly revises and adds to what should really follow from his official definitions. In so doing he involves himself in contradictions and yet arrives at true thoughts, which, one must admit, come into his mind in a purely haphazard way. His sentences express true thoughts, if they are rightly understood. But if one tried to understand them in accordance with his own definitions, one would go astray. (*PW* 222/ *NS* 239–240, 'Logik in der Mathematik' (1914))

Here Frege issues a sharp statement of the point that the best mathematicians may fail to have a clear grasp of senses or thoughts they think with. Nevertheless, he emphasizes that the expressions such mathematicians use express definite senses, definite thoughts (*Gedanken*). The presupposition of the whole passage is that the science of arithmetic and the words that Weierstrass uses already express the relevant senses and denote the relevant concepts. The lack of clarity or definiteness resides purely in the person: it is 'psychological' not 'logical' or 'arithmetical'. The mathematician is thinking with the sense, and expressing it in his writing, but not grasping it clearly.[13]

Frege articulates the same view, in a passage we quoted earlier (*PW* 211/ *NS* 228, 'Logik in der Mathematik' (1914)), at the conclusion of a long and interesting discussion of philosophical or foundational analyses—analyses of the sort that he had pursued in trying to establish logicism. The suggestion is that insofar as an analysis like the one he attempted in *Foundations of Arithmetic* is true but still subject to doubt, the difficulty resides in the person's inability to grasp the sense clearly.

These last two passages from 1914 contain the resolution of the apparent tension between the passages that suggest that grasping the sense of an expression is relatively easy and those that indicate that it is quite difficult. It is possible to express a sense and think by 'grasping' senses even though one lacks a 'clear' or 'thorough' 'grasp' of the sense. The sort of grasping that is necessary for successful communication and ordinary thinking is different from the sort (or level) of grasping necessary to articulate the senses of one's expressions through other terms.

[13] The remark comparing Weierstrass's plight with that of the ordinary mathematician when he thinks with the notion of the integral is useful in suggesting that thinking can go on without a clear, thorough grasp of the sense (the thought). On the other hand, it is important to recognize that although the notion of the integral has a fuller analysis that was coming to be grasped by the mathematical community, the notion of number that Weierstrass has not fully grapsed (and hence, presumably even the notion of the integral!) was not, in Frege's view, fully understood by any mathematician prior to Frege's own work. So the remarks that he applies to Weierstrass are applicable to the whole mathematical community.

In one respect, Frage's distinction between levels of understanding is quite ordinary. We commonly attribute thoughts and statements to people on the basis of their linguistic expressions even though those people cannot produce acceptable definitions or explications of the terms they use. That is, we interpret our attributions in standard ways even though the people to whom we are making the attributions are unable to articulate the standard meanings of the expressions they use. This practice extends even to cases where the people in question misunderstand the relevant expressions.

The striking element in Frege's view is his application of this distinction to cases where *the most competent speakers, and indeed the community taken collectively*, could not, even on extended ordinary reflection, articulate the 'standard senses' of the terms. The view is that the most competent speakers may be in the same situation as the less competent ones in expressing and thinking definite senses which they cannot correctly explicate or articulate. Definite senses are expressed and 'grasped' (with merely the weak implication that they are thought with), even though no one may be capable of articulating or explicating those senses (grasping them clearly and analytically).

Frege does not introduce a distinction between sorts or levels of sense-expression that parallels his distinction between sorts or levels of understanding or 'grasping'. Here it is clear that sense-expression is not seen as supervenient on use and individual or communal capacities in the way that conventional (or idiolectic) linguistic meaning is commonly taken to be.

As I indicated earlier, I think that Frege's conception attempts to bridge the gap between actual understanding and actual sense-expression by means of a normative concept—that of the deeper foundation or justification for actual understanding and usage. It would be incorrect to see Frege's notion of sense-expression as separated from that of actual understanding and use. But it does depend on the possibility of a projection beyond actual understanding and use.

Even the ordinary notion of conventional linguistic meaning depends on a projection. Actual usage is interpreted in terms of a standard that is drawn from such usage. But the standard depends on a certain sort of rationalized ordering of that usage. Think of how dictionary definitions are arrived at. Frege's idealization is, however, much more radical. His conception of an 'ideally competent speaker' is much further removed from that of actual competence than is the conception of an ideally competent speaker that is current in discussions of conventional (or idiolectic) linguistic meaning.

Frege's ideal speaker is not simply an ordered composite of what is—or would be, under certain fairly ordinary conditions—recognized to be the best extant usage and understanding. Frege's ideal speaker may have an understanding that is fundamentally better than, even substantially different from, anything anyone has yet achieved. Understanding in the relevant sense is seen as an achievement that brings with it insight into substantive truths. In this, Frege is at one with the leading representatives of the rationalist tradition. (Cf. *FA* p. vii, *PW* 12–13, 222/*NS* 16–17, 'Booles rechnende Logik und die Begriffsschrift' (1880–1881); *NS* 240, 'Logik in der Mathematik' (1914).) As with Plato and Descartes, deep understanding of one's thoughts—and of the senses (forms, ideas) one thinks with—is not separable from the deepest sort of knowledge (*KS* 122–124, 'Über das Trägheitsgesetz', O 157–161; *KS* 369, 'Die Verneinung', O 150; *PW* 33/*NS* 37, 'Booles rechnende Logik und die Begriffsschrift' (1880–1881)).

Analyses that articulate the senses of expressions do not constitute degenerate knowledge for Frege, any more than logic and arithmetic themselves do. Nor are they true 'purely in virtue of sense'. Analytical insight into the nature of one's senses or thoughts may involve a legitimate feeling of having been 'committed' to the relevant knowledge all along—may produce some sense of reason's recalling its origins. But from the point of view of what informs one's explications, one's actions, and one's operative understanding, such knowledge may count as new and substantive. It is not *merely* reflection on one's thoughts or meanings in the sense that the empiricist-positivist tradition in this century would represent it. This aspect of Frege's view has been seriously neglected because it is so out of keeping with the meaning-is-use approaches that stretch from the Vienna Circle to the present.

What is original about Frege's interpretation of the relation between conceptual insight and the deepest sort of propositional knowledge is his reversal of the traditional order of priority, and his emphasis on the role of theory in attaining such knowledge. Traditionally, among some rationalists as well as empiricists, conceptual mastery was considered a precondition for judgement. And such mastery was interpreted in terms of what were presumed to be non-conceptual abilities, such as vision. This model is almost completely absent in Frege's work. There are, as I noted earlier, significant traces in Frege of the pre-Cartesian, Platonic picture of *a priori* knowledge as outer vision: vision of an eternal reality that informs and constitutes one's reason. But the traditional picture does not inform Frege's work in the traditional way. His model is not vision but theory.

Kant preceded Frege in insisting on the priority of judgement over non-propositional cognitive capacities. But Frege went beyond this Kantian view in two ways. In the first place, he developed a logical theory that demonstrated the fruitfulness of an account of judgement in analysing non-propositional parts and non-propositional cognitive capacities.[14] That is, Frege gave deep and detailed grounds for accepting the order of priority that Kant announced.

In the second place, Frege held that the analysis of judgement—and more generally, logic—is inseparable from *theoretical activity*. Unlike Kant, Frege

[14] This point is developed at considerable length in my 'Frege on Truth' (Ch. 3 above), especially the first and last sections.

was an original logician. Perhaps because of this he was more sensitive to the difficulties of arriving at a satisfactory logical theory and to the possibility that reflection on logical matters, as well as in other domains, has to be tested in a comprehensive theoretical context. For Frege, logic has to be discovered. It is not transparent to reflection. It has to be discovered through checking proposals as applied to discriminating reflection on ordinary theoretical usage in logic, mathematics and other domains. Frege thought that one could be sure of attaining thorough grasp of the senses of one's terms only through understanding their logical roles and their proper roles in a good theory. Such roles were guaranteed to be perspicuously revealed only in a logically perfect language that was ideally adequate to its subject matter.

Frege is a traditional rationalist in his belief in *a priori* knowledge, in his association of the deepest substantive knowledge with conceptual mastery, and in his modelling philosophy itself on logic and mathematics. But he goes beyond the tradition: Even *a priori* knowledge and non-propositional understanding (or conceptual mastery) were construed as grounded in logical analyses of theory, and checked against good theoretical usage. What makes Frege modern and entirely original is his combination of his traditional views about apriority, understanding and philosophy with a pragmatic understanding of the epistemology of logic. Thus Frege held a fallibilist and theory-based account of *a priori* knowledge, understanding and philosophical inquiry.[15]

Frege expresses his pragmatic perspective on understanding and *a priori* knowledge only occasionally. But it appears at all periods of his career. For example, in an unpublished manuscript from 1880–1, he writes:

> All these concepts have been developed in science and have proved their fruitfulness. For this reason what we may discover in them has a far higher claim on our attention than anything that our everyday trains of thought might offer. For fruitfulness is the acid test of concepts, and scientific workshops the true field of study for logic. (*PW* 33/ *NS* 37, "Booles rechnende Logik und die Begriffsschrift" (1880–1881))

The same point is developed at somewhat greater length in a passage of 'About the Law of Inertia' (1891). There Frege holds that thorough grasping of concepts has resulted to a large extent from recognition that the accepted delineation (*Begrenzung*) is 'blurred, uncertain, or not the one that was sought' (*KS* 123, O 159). He develops the idea (*KS* 122–124, O 157–161) that only through the development of scientific theory, whether in physics, chemistry or logic, does one achieve a thorough grasp of one's concepts. Similar remarks occur in the later period. (Cf. 'Die Verneinung', *KS* 369, O 150; *BL* 7; 25/ *GG* i. pp. x, xxvi.)

[15] This point makes it appropriate to say that for Frege philosophy is modelled on and continuous with science. But the relevant sciences are logic or mathematics, not natural science as it is for the positivists, Carnap and Quine. In this view, Frege is again a traditionalist (emulating Plato, Descartes, Leibniz, and Kant.)

Partly because of Frege's work, and the work of those such as Wittgenstein, Carnap and Quine who were influenced by it, the connection between understanding and theory has become a commonplace. What is still unassimilated, in my opinion, is the way Frege attaches ideal understanding and actual sense expression to ideal theory. It is his explication of sense-expression in terms of theorizing beyond any that may have actually been carried out that sets his conception of sense apart from more ordinary conceptions of conventional (or idiolectic) linguistic meaning.

VII

I wish now to turn briefly to the first of the two presuppositions behind Frege's view that the ultimate foundation and justification of current linguistic practice may supplement ordinary understanding in such a way as to attach it to a definite sense that no one may currently be able to articulate adequately, or thoroughly grasp. The presupposition is that some of our practices are founded on a deeper rationale or on deeper aspects of 'reality' than anyone may have presently understood.

This presupposition is, of course, bound up with deep and contentious philosophical issues. Systematic development of the idea is anything but trivial. Although I think that the idea need not be developed in foundational terms, Frege clearly regarded his senses, at least in some contexts, in the light of some final state of cognitive achievement.

One naturally thinks here of some Peircean conception. Although like Peirce Frege had a notion of ideal science, he did not construe it in terms of agreement, or in terms of any epistemological notion (such as the application of best canons of reason to all possible evidence). Such accounts would not express the notion of an ideal science for Frege because he thought that there is no conceptual guarantee that any human agreement, or any humanly applied epistemic methods, would not be mistaken. He thought it conceivable that everyone might agree on what is not correct. Where a theory is incorrect, there would always remain room for a better theory that might provide better understanding of the senses of terms in the original theory.

There are, of course, epistemically oriented accounts of ideal science that recognize these limitations and try to idealize beyond them: we should not limit the account to *human* agreement or observational capacities, or to what is available to *us* as possible evidence, or to *our* inductive methods. All these might be improved upon. But when one projects beyond human practices, the idealizations become much less clear. Moreover, they become subject to the suspicion that they are covertly relying on the assumption that they are practices that are ideal in the sense that they yield a *true theory*. Such reliance would make epistemic notions presuppose semantical notions, giving up the spirit of Peirce's idea. But short of reliance on semantical notions, it is unclear how such epistemic explications can give a satisfactory account of the notion of ideal science—and hence of Frege's notion of sense. In any case, for Frege, the notion of truth is not to be reduced to epistemic terms.

For Frege, thoughts, the senses of declarative sentences, are conditions on truth. Conditions on truth are best understood, and senses are best articulated, from the standpoint of a comprehensive true theory. This is not to say, of course, that only expressions in a true theory have sense. The forerunners of such a theory also express senses, even if the senses are not thoroughly grasped. Moreover, non-denoting terms express senses. So even theories that are on the wrong track and have no justifying rationales express senses, possible ways of thinking. The point is rather that adequate explication of the senses of expressions is *guaranteed* only in a theory that cannot suffer further improvement or correction.

Frege did not attempt the reduction in reverse. He did not try to explicate epistemic notions by means of his notion of truth. He was as aware as anyone that the notion of truth does not provide an epistemic touchstone for judging the credibility of theories (cf. the opening pages of 'The Thought'). Thus he seems committed to the view that one has no epistemic guarantee, beyond all conceivable doubt, that one has arrived at a true theory.[16] But one is guaranteed an understanding of one's senses only from the point of view of an 'ideal science'—a true theory. So how to explicate the senses of one's terms is in principle just as subject to theoretical debate as to whether one's theory is true. The terms and methods of debate are surely different. But there is no difference as regards an epistemic guarantee that one is right.

The notion of truth provides a normative ideal that (in inevitable conjunction with concrete epistemic practices) enables us to conceive of terms' having senses that are not fully explicable in terms of usage and understanding. An ideally competent speaker is one who can give ideally thorough explications

[16] Frege does regard the basic truths of logic as self-evident. But he allowed that one might be mistaken about what is self-evident, if one has a less than thorough grasp of the thoughts involved. And the question of whether one has a thorough grasp is answerable by reference to the success of the theory in which the thoughts are embedded, cf. my 'Frege on Extensions of Concepts' (Ch. 7 below), pp. 30–4 (pp. 295–298, this volume).

It is worth pointing out here that the 'final' explications of an expression are not, at least are not in general, expressions of the same sense. Frequently they are part of fixing the sense of the expression being explicated. Frege thought that only in the context of a proposition (and more generally, a theory) could the sense of scientific expressions be fully understood. The senses of expressions could be fully grasped only by grasping equivalences given by ideal scientific explications, or by otherwise understanding the contribution of those expressions to a theory. Of course, since most of the envisioned ideal explications would come as discoveries, it is possible to doubt them (even if the doubt depends on less than full analytic mastery of the senses), while not doubting the corresponding self-identities. So by Frege's test for the identity of senses, the senses of the *explicans* and *explicandum* would be different. Of course, many expressions will express senses but will not appear in ideal science at all (e.g. denotationless expressions). The point is not that senses are expressed only in an ideal science, and certainly not that every term is incompletely understood. The point is rather that terms in science may not be fully understood, and that full understanding is guaranteed only when one has an ideal science.

and has an ideally thorough grasp of the senses of the expressions of his language. Having tied adequate explication and thorough grasp to having fundamental knowledge, and having tied fundamental knowledge to common theoretical activities (rather than indubitable insight), Frege must see full understanding as guaranteed only by completely fundamental and completely satisfactory (true) theory. It is this combination of a rationalist notion of understanding with a pragmatic epistemology that makes Frege's view unique and intriguing.

Since Frege did not develop this conception, it is important to reflect on it—and be prepared to develop it—in a flexible way. The view is most immediately attractive as applied to terms in mathematics and the empirical sciences. It is less plausible, at least initially, with respect to ordinary expressions that have no role in the sciences. The idea that one might obtain some significant improvement in our understanding of them may seem rather far-fetched. And surely, one is inclined to think, there are some expressions in ordinary discourse for whose understanding we would simply refuse to recognize any further authority than communal practice.

It is noteworthy that Frege always applies his remarks about the most competent users' incomplete understanding of their own expressions to terms that have a role in a scientific enterprise. Perhaps he would have given a different theory for non-scientific terms that are not grasped sharply. He may have held that such terms are associated with a cluster of senses. Where none of the senses is grasped sharply, no determinate thought is entertained. Or (what is more suggested by his actual remarks about vague terms) a definite sense might be assigned in the context according to how the thinker would draw a line if pressed. Perhaps he would have seen no need, in the case of ordinary non-scientific terms, for the possibility that a particular, unforeseen sharpening might be the correct one. Surely this possibility is sometimes closed.

On the other hand, the possibilities for rational improvement that extends elements already present in a practice are difficult to circumscribe. What underlies the attractiveness of Frege's conception, to my mind, is not some prior view of what a science is. It is a combination of two more general views. One is the view that full understanding of our thoughts and senses is not independent of knowledge of the subject matters about which we think. The other is that neither individually nor communally do we have infallible access to the truth about those subject matters. These two views jointly press one toward Frege's conception. And they retain some force even as applied to thoughts outside the systematic sciences.[17]

Let us return briefly to our initial puzzle about vagueness. The puzzle helped signal the distinctive character of Frege's conception of sense. The

[17] For a beginning at developing these ideas, see my 'Intellectual Norms and Foundations of Mind', *The Journal of Philosophy*, 83 (1986), 697–720.

account of sense that we have given helps, in turn, to illumine Frege's otherwise odd views about vagueness. Through theory one arrives at the sharpening of the application of expressions. This sharpening enables one to go beyond conventional or idiolectic linguistic meaning to arrive at a full understanding of one's expressions, a thorough grasp of their senses. Until a powerful theory is developed and logical analysis of its language is achieved, a full grasp of the senses of one's own expressions may be lacking.

Frege's conception of sense as deeply connected to that of a true theory also sheds light on, though it may not justify, his view that no senses have vague boundaries. Where senses are ways of thinking ideally or purportedly appropriate to a true theory, they are ways of thinking that ideally or purportedly reflect reality. Frege believed that vagueness could not infect reality itself—the objective entities which our thoughts denote. I suspect that only if one explicates reality in terms of mind or meaning does the notion of vagueness in reality make any sense. And Frege resisted such explications, partly out of allegiance to common sense, partly because of the sort of considerations that we have been rehearsing. So he concluded that senses, ways of thinking about reality, could not be vague. Only our grasp of senses could be.

This line of thought is, of course, more attractive for ways of thinking that are ideally fitted to reality than for ways of thinking that are merely purportedly fitted to reality. Some senses, as we noted, do not fit into an ideal science. They would not be expressed by a true theory. It is unclear how such an argument that senses cannot be vague could be applied to such senses. Perhaps the old charge that Frege's idealizations are too imperial still sticks. But it seems to me that his motivations are far more subtle and interesting than the traditional criticisms of his views on vagueness recognize.

I do not want to suggest that these issues are anything like as simple as I have been representing them. My primary point has been to try to indicate how Frege's odd-sounding views about vagueness are motivated by his realist conception of the world and his rationalist conception of sense.

VIII

I will close by noting in an extremely sketchy way some points about the subsequent historical fate of Frege's notion of sense.

Russell's views of understanding and knowledge began the historical process of obfuscation. From the beginning, Russell allowed little or no room for the idea that one could think with notions that one only partially understands. Russell assimilated understanding to a non-propositional, vision-like conception of knowledge, called acquaintance, which propositional abilities were supposed to presuppose. Much of the motivation for his semantical work lay in his notorious 'principle of acquaintance', which antedates

'On Denoting' *Mind*, 14 (1905), 479–493. According to this principle one has to be 'acquainted with' every constituent of a 'proposition' that one thinks. The key point for us is that acquaintance was understood in terms of an infallible, direct and absolutely complete mastery of the propositional constituents. Eventually, Russell held that the only things with which we can be acquainted are universals, present sense-data, and our selves. This approach led, of course, to much discussed difficulties with the theory of reference to individuals. But it did not allow serious consideration of the understanding of universals.

Encouraged by his interest in establishing logicism, Russell did emphasize that through analysis one could enlarge one's powers of understanding and gain insight into the foundations of notions that one is already thinking with. But because he treated understanding in terms of acquaintance rather than in terms of the use of a theory, Russell never allowed this emphasis to threaten the principle of acquaintance and the view of understanding and propositional thought that rests upon it. Thus the implicit tension between the principle of acquaintance and the natural view that with 'analysis' we obtain a better understanding of the analysanda never comes to a head in Russell's work.

Russell's communication of Frege's views to the rest of the philosophical world did nothing to call attention to the difference between his own views and Frege's on these points.

It is not clear to me to what extent the Vienna Circle reacted against Russell's vulnerable conception of understanding and to what extent it simply bypassed it. What is clear is that the Vienna Circle struck off in a very different direction. The early, more phenomenalist stage of the movement bears comparison to Russell's work in its interest in explicating understanding in terms of a construction from basic, rather specially certain types of cognition (the having of phenomenal experiences). But the prominence of this view diminished as the movement developed. Moreover, from the beginning the movement was much more ambitious about giving a detailed account of understanding than Russell was.

There seem to me to be two key ideas in the positivist programme that are relevant to our theme. One is the view that what there is to be understood, 'meaning', is to be reduced to actual procedures that express or constitute actual understanding. That is, meaning is identified with cognitive or theoretical usage. The other key idea is that both meaning and understanding are to be accounted for in terms of verification procedures.

The standard histories of the period concentrate on the second idea. But it seems to me that the first is equally momentous. Roughly speaking, it contains two moves: the reduction of what is understood to actual understanding, and the explication of actual understanding in terms of presently articulateable abilities or experiences. These moves seem to me to be fundamental, but largely unexamined in the work growing out of the Vienna Circle movement.

The influence of these moves is evident in the explications of Frege that have been taken to be authoritative until recently. Carnap's explication of Fregean sense in *Meaning and Necessity* is a prime example. He writes, 'The concepts of sense and of intension refer to meaning in a strict sense, as that which is grasped when we understand an expression without knowing the facts.'[18] Quine takes up the same line, with fewer positivist commitments, when he explains Frege's conception, without comment, in terms of linguistic meaning. Most (though not all) of Kripke's criticisms of Frege depend essentially on identifying sense with an ordinary conception of linguistic meaning. Dummett's account of Frege's notion of sense, though much more nuanced and in many respects more insightful, takes up the same theme—again influenced by preconceptions about the relation between 'meaning' and 'use'.[19]

What made it historically difficult to appreciate the difference between Frege's conception of sense and modern notions of linguistic meaning was the domineering role of the idea that originated with the Vienna Circle: the reduction of what is understood to actual understanding and the explication of understanding in terms of actual articulateable abilities or experiences. The conception of meaning as used in the later Wittgenstein, the attempt to reduce philosophical problems to questions about ordinary language, and Quine's holistic liberalizations of verificationist conceptions of meaning, are all marked by these moves.

Parallel doctrines in the philosophy of mind are equally founded on the reduction of what is thought to what the thinker can presently do. Behaviourist and functionalist theories of mental content, Wittgenstein's puzzles about the possibility of following a rule, and Quine's claims that mind and meaning are fundamentally indeterminate—all depend on the idea that what is understood is to be reduced to or accounted for purely in terms of present abilities (where these are described in such a way as not to make explicit use of intentional notions like understanding or thought).

There is, of course, considerable force—not to say epistemic comfort—in this idea. But its dominance has engendered a certain blindness to the character and, I think, power of the very different point of view that motivated Frege. Appreciating and developing that point of view may help illumine some of the dark dead-ends of twentieth-century philosophy.

[18] Rudolf Carnap, *Meaning and Necessity*, 2nd edn. (Chicago: University of Chicago Press, 1956), 125. Carnap goes on to identify sense with intension, so that logically equivalent terms have the same sense. He explains logical equivalence in modal terms, thus further obscuring Frege's view. The explication of 'sense' in modal rather than cognitive terms continued in discussions of Frege into the 1970s and recurs occasionally even now. Such explication of course has virtually nothing to do with Frege's original conception and his motivating problems.

[19] W. V. Quine, 'Two Dogmas of Empiricism', in *From a Logical Point of View*, 2nd edn. (New York: Harper Torchbooks, 1961), 21; Saul Kripke, *Naming and Necessity* (Cambridge, Mass.: Harvard University Press, 1980), e.g. 53–4; Michael Dummett, *Frege: Philosophy of Language* and *The Interpretation of Frege's Philosophy*. Some of these same points are discussed in my 'Sinning Against Frege' (Ch. 5 above).

This page intentionally left blank

Part III

Rationalism

This page intentionally left blank

7 *Frege on Extensions of Concepts, From 1884 to 1903 (1984)*

Russell's paradox defeated Frege's attempt to demonstrate logicism—the view that the mathematics of number can be derived from the axioms of logic together with definitions, in logical terms, of mathematically primitive expressions. The paradox indicated that Frege's logic is inconsistent and that the notion in terms of which he tried to define the cardinal numbers is defective. The defective notion, that of the extension of a concept, has remained interesting because it is motivated by intuitions that, arguably, play an ineliminable but inadequately understood role in modern set theories.[1] The development of Frege's views is also historically interesting. Frege was uncertain about the crucial notion and the paradox-producing axiom from the outset. He experimented with alternatives. What can be pieced together about his reasoning suggests deep tensions in his thought on fundamental matters.

This paper is purely historical. It concentrates on the period from the publication of *The Foundations of Arithmetic* in 1884 to the publication of *The Basic Laws of Arithmetic* (first volume, 1893; second, 1903). I shall trace the development during this period of Frege's notion of the extension of a concept.

I. CONCEPTUAL UNCERTAINTY AND FREGE'S RATIONALISM

The story begins with two remarkable passages in *The Foundations of Arithmetic*. Both are enigmatic. Both appear to express a lack of commitment to the notion of the extension of a concept (*Umfang eines Begriffes*, or

I am indebted to the referee of *The Philosophical Review* for suggestions that led to improvements.

[1] Cf. Charles Parsons, "Some Remarks on Frege's Conception of Extension", in M. Schirn (ed.), *Studien Zu Frege* (Stuttgart: Frommen-Holzboog, 1976); "Sets and Classes", *Noûs*, 8 (1974), 1–12; "What is the Iterative Conception of Set?", in R. E. Butts and J. Hintikka (eds.), *Logic, Foundations of Mathematics, and Compatibility Theory* (Dordrecht: D. Reidel, 1977). The first article discusses Frege's notion of extension in the period from 1902, the discovery of the paradox, to 1925, Frege's death.

Begriffsumfang). The first occurs in a footnote to the passage in which Frege first defines the expression 'the Number that belongs to the concept F'. He defines the expression as: the extension of the concept *numerically equivalent to the concept F*. This is Frege's first use of the notion of the extension of a concept. And he immediately attaches the following famous footnote:

(A) I believe that for "extension of the concept" we could write simply "concept." But two objections could be raised:

1. that this contradicts my previous statement that the individual numbers are objects, as is indicated by means of the definite article in expressions like "the number two" and by the impossibility of speaking of ones, twos, etc. in the plural, and also by the fact that the number makes up only an element in the predicate of a statement of number.
2. that concepts can have the same extensions, without coinciding (*zusammenfallen*).

I am however of the opinion that both objections can be refuted; but here that would lead us too far afield. I assume that it is known what the extension of a concept is. (*FA*, 80 n)[2]

The second striking passage occurs in the summation of the book. After reviewing the difficulty of providing an explicit definition for numerical expressions and citing the above-mentioned definition, Frege writes:

(B) In this we take for granted the sense of the expression "extension of the concept." This way of overcoming the difficulty will not win universal applause, and many will prefer to remove the doubt in question in another way. I attach no decisive importance to bringing in the extension of a concept. (*FA*, 117)

Several commentators have called attention to these passages. Most simply note the fact that Frege is uncharacteristically obscure and indefinite, especially given the fundamental importance of the definition in question. Some have misread the passages. I think that we can come to understand them by giving them a context and by reflecting on other writings that make reference to them.

[2] Textual references occur in the text. The following works of Frege will be cited with the abbreviations noted: *The Basic Laws of Arithmetic* (*BL*), trans., Furth (Berkeley: University of California Press, 1967); *Begriffsschrift (B)*, ed., Angelelli (Hildesheim: Georg Olms, 1964); *The Foundations of Arithmetic* (*FA*), trans., Austin (Evanston, Ill.: Northwestern University Press, 1968), (the pagination in the translation is the same as that in the original German); *Die Grundgesetze der Arithmetik* (*GG*) (Hildesheim: Georg Olms, 1962); *Kleine Schriften* (*KS*), ed., Angelelli (Hildesheim: Georg Olms, 1967); *Nachgelassene Schriften* (*NS*), ed. Hermes Kambartel, Kaulbach (Hamburg: Felix Meiner, 1969); *Philosophical and Mathematical Correspondence* (*PMC*), trans., Kaal (Chicago: University of Chicago Press, 1980); *Posthumous Writings* (*PW*), trans., Long and White, (Chicago: University of Chicago Press, 1979); *Translations from the Philosophical Writings of Gottlob Frege* (G&B), ed. Geach and Black (Oxford: Basil Blackwell, 1966); *Wissenschaftlicher Briefwechsel* (*WB*), ed. Hermes *et al*. (Hamburg: Felix Meiner, 1976). Translations are mine. Changes from the excellent translations by Austin, Furth, and others are usually minor, though especially in Austin's case, they often affect the tone and are more literal.

The first question that arises concerns the nature of Frege's ambivalence. Is Frege indicating that he thinks that there are alternative, significantly different, but equally good definitions within logic of 'cardinal number'?[3] I think that the answer to this question is "no." Rather, Frege was uncertain both about his whole approach to the problem through explicit definition (B), and, granted that approach, about the use and explication of 'extension of a concept' (A). The full evidence for this answer must emerge slowly.

An initial and weighty consideration against the view that in (A) and (B) Frege was indicating that there are equally good non-equivalent ways of defining the numbers within his logic is simply a matter of common sense. Frege was starved for recognition after the poor reception of his *Begriffsschrift* in 1879. If he had had in mind fundamentally different non-equivalent ways of defining number he would have announced the view, displayed the definitions, and explained the alternatives in some detail. In fact, he suggests one definition, but is lame in his remarks about it. The posture contrasts sharply with the confident, almost overbearing, tone of the rest of the book. It suggests uncertainty, not confidence in a plurality of ways of reaching his goal.

A larger reason for taking (A) and (B) to evince uncertainty derives from considering Frege's project from Frege's perspective. Frege regarded the *Sinn* and *Bedeutung* of 'Number' as not completely determined by conventional mathematical usage, at the time of his attempt to define it in his logic.[4] That is, the conventional significance of the term did not fix its *Sinn* or *Bedeutung*.

Sense and conventional significance diverge elsewhere in Frege's philosophy—in his views on indexicality, for example.[5] The reasons for the divergence in this case are that mathematical language is logically or structurally unperspicuous, and that mathematical usage is, by Frege's standards, vague. For our purposes, vagueness is the key problem. To be fully competent by conventional standards one did not have to grasp something that settled all the questions about the application of the term that a genuine sense had to settle.

Thus, from Frege's point of view, "grasp" of the conventional, mathematical significance of 'Number' might be "complete" (one might count as fully understanding its current usage), while one's grasp of the term's sense was incomplete and uncertain. Frege held that senses (or "concepts"—which are

[3] An affirmative answer is proposed by Paul Benacerraf in "Frege: The Last Logicist", *Midwest Studies*, 6 (1981), 29 ff. Benacerraf claims that Frege's account of number was much more similar to modern accounts, in which alternative definitions of numbers within set theory are allowed, than is usually thought. He claims that (A) and (B) indicate that Frege did not expect even reference to be preserved by his definitions. Benacerraf's argument partly rests on a further passage in *The Basic Laws of Arithmetic* (sec. 10). I discuss this passage in "Frege on Truth" (Ch. 3 above).

[4] Frege had not distinguished *Sinn* and *Bedeutung* in 1884, but the point would apply to both. The term 'Begriff' ('concept') had also not acquired the technical use it later acquired in Frege's work.

[5] Cf. my "Sinning Against Frege", *The Philosophical Review*, 88, no. 3 (1979), 398–432 (Ch. 5 above).

in his earlier work an amalgam of sense and denotation, and later just denotation) are not themselves vague or indefinite. He held them to be eternal and absolutely determinate in their corresponding *Bedeutungen* (or in the objects that fall under them). The vagueness problem lay with human understanding and with conventional usage, not with the senses, or concepts, or their expression.

Strictly speaking, from Frege's point of view, no one had fully grasped and mastered the concept of number or the *sense* (as distinguished from the conventional significance) of 'Number' by the time he wrote *Foundations* in 1884. By scrutinizing passages (A) and (B), I think we will find that Frege would have had to apply this point even to himself.

I will not have the space to justify this interpretation in depth. But I will discuss some of the passages that support it. In *Begriffsschrift* section 27 Frege indicates that a vague expression does not express something "judgeable." This sort of remark is faintly echoed only once later in his career (*PMC* 114–115/*WB* 183, Frege to Peano, 9/29/1896). But he frequently states that vague concept expressions lack *Bedeutung* (e.g. *PW* 122, 155, 179/*NS* 133, "Ausführungen über Sinn und Bedeutung" (1892–1895); *NS* 168, "Begründung meiner strengeren Grundsätze des Definierens" (1897/98 oder kurz danach); *NS* 193–194, "Über Schoenflies: die logischen Paradoxien der Mengenlehre" (1906); *PMC* 114/*WB* 183, Frege to Peano, 9/29/1896; *KS* 122–123, "Über das Trägheitsgesetz", O 158–160). Now the sense for 'Number' that Frege sought to define would certainly fix a definite *Bedeutung*. So if 'Number' in its conventional usage is vague, its conventional understanding and explication will not articulate the concept denoted by the term, or the sense that it expresses (cf. *GG* ii., secs. 69–70).

In several passages Frege implies that the conventional usage of such expressions as 'Number' *is* vague—not merely unperspicuous. At the end of *Foundations* (*FA* 110), he indicates that the "meaning" (*Bedeutung*) (conventionally) assigned to 'sum' and 'product' before the introduction or discovery of complex numbers said nothing about the admissibility or inadmissibility of certain further assignments of "meaning." (Cf. *PMC* 63/*WB* 96, Frege to Husserl, 5/24/1891; G&B 28/*KS* 131, "Funktion und Begriff", O 12–13.) Nothing in then current mathematical usage determined how to extend the conventional "meaning" of 'Number' to cover cases about which it was previously mum. Further philosophical and mathematical considerations had to be appealed to.

Foundations can be read as part of an attempt to produce the final in a series of sharpenings of the "meaning" of 'Number'. Referring, in the introduction, to the task of understanding the concept of number, Frege writes,

What is known as the history of concepts is a history either of our knowledge of concepts or of the meanings of words. Often it is only through great intellectual labor, which can continue over centuries, that a concept is known in its purity, and stripped of foreign covering that hid it from the eye of the intellect. (*FA* p. vii)

Seven years later, Frege embroiders this passage:

> For the logical concept, there is no development, no history. . . . If instead of [this sort of talk] one said "history of the attempt to grasp a concept" or "history of the grasp of a concept," it would seem to me far more appropriate; for the concept is something objective that we do not form and is not formed in us, but that we try to grasp and finally, it is hoped, really grasp—if we have not mistakenly sought something where there is nothing (*KS* 122, "Über das Trägheitsgesetz", O 158).

Clothing or covering is Frege's metaphor for a concept's (or a sense's) verbal expression. The implication is that mastering conventional mathematical usage of the term 'Number' does not suffice for grasping the logical concept that his labor seeks to reveal. Note also the implication in *FA* p. vii that there is a single "pure" concept to be found once one's intellectual labor is rewarded. Both passages highlight the difficulty of grasping a logical concept.

Section one of *Foundations* suggests that mathematical concepts such as function, continuity, limit, infinity stand in need of sharper determination. (Lack of sharpness is cognate for Frege with vagueness.) The ideal of "sharp grasp of concepts" is said to be the fundamental aim of the book. (Austin's translation obscures this point slightly by translating "scharf zu fassen" as "sharp definition.") Section 2 indicates that the concept of number is Frege's prime target in the attempt to attain "sharp grasp." In the light of these passages, the remarks about fixing a "sharply bounded" concept of number in *FA* 74, 79 should be seen not only as remarks about the particular stage of Frege's own inquiry, but also about the status of conventional understanding in mathematical science. In *Foundations* 81, he acknowledges that the remark that one number is wider than another, though supported by his definitions, is neither supported nor opposed by ordinary usage (*FA* 81). As noted, the last pages of the book return to the vagueness theme. The discussion of complex numbers provides a model for his own method (*FA* 110 ff.): philosophical and new mathematical considerations must supplement conventional usage.

Articulation of the commonness of incomplete understanding recurs in Frege's discussion of "analytic" definitions in an unpublished work from 1914. I know of nothing that prevents seeing his remarks there as congenial with his earlier views. He says that in analyzing the sense of an expression with long-established usage, one may be uncertain whether the *analysans* gives the sense of the older expression. In such cases, he says, one can drop the older sign from the formal system and introduce a new counterpart whose sense is then "stipulatively" defined within the formal system. He continues:

> (C) How is it possible, one can ask, that it be doubtful whether a simple sign has the same sense as a complex expression, if not only the sense of the simple

> sign is known, but also the sense of the complex expression can be recognized from the way it is put together? In fact, if the sense of the simple sign is really clearly grasped, then it can not be doubtful whether it agrees with the sense of the complex expression. If this is open to question even though the sense of the complex expression can be clearly recognized from the way it is put together, the reason must lie in the fact that the sense of the simple sign is not clearly grasped, but appears as through a mist only with blurred outlines. (*PW* 210–211/*NS* 227–228, "Logik in der Mathematik" (1914). Cf. also *PW* 207/*NS* 224.)

The mist metaphor is closely associated with the metaphors of clothing or veiling, which are used to indicate that ordinary linguistic usage sometimes prevents a clear grasp of the associated concepts or senses (*FA* p. vii). The reason for an incomplete grasp of a mathematical concept or sense, in Frege's view, lies in the logical insufficiencies and vagueness of ordinary mathematical usage. These drawbacks are inseparable, in Frege's view, from insufficient foundational knowledge.

The project of defining 'Number' in purely logical terms cannot then be seen as attempting to specify its significance in any ordinary sense of 'significance'. For the sense and *Bedeutung* of 'Number' are not its conventional significance, and are only constrained—not fixed—by standard mathematical understanding and usage. On the other hand, arriving at a definition of 'Number' is not a matter of stipulation (except as seen within the confines of the formal system). Mathematical practice places substantial restrictions on a good definition. Moreover, I think that any room for arbitrariness in attaching word to concept, or to sense, that survives these restrictions was conceived by Frege as arbitrariness relative only to conventional, mathematical usage. Frege was not interested in making just any mathematical model of arithmetic. The very project is foreign to his methodology. I think that he sought a single objective conception of number that underlies and fully rationalizes mathematical practice (cf. my "Frege on Truth" (Chapter 3 above)).

These considerations raise a question about the nature of sense-expression (and concept denotation). How could the term 'Number' indicate a definite "concept" when all current mathematical understanding and usage failed to determine a sense or concept? Even granted that senses and concepts themselves are independent of human minds or human activity, how can Frege regard sense-*expression* (or the denotation of concepts by language) as even partly independent?

There is scattered evidence that Frege was aware of the force of these questions. For example, as we have noted, in *Begriffsschrift* section 27 he writes of a particular vague expression that it does not have a judgeable content. In 1880–81 he writes that the logical relations implicit in a particular logically unperspicuous expression *are not expressed*, but must be guessed at (*PW* 13/*NS* 14, "Booles rechnende Logik und die Begriffsschrift"

(1880–1881)). In a letter to Peano in 1896 he seems to indicate that vague expressions lack sense as well as *Bedeutung* (*PMC* 114–115/*WB* 183, Frege to Peano, 9/29/1896). And Frege repeatedly states that vague concept words lack *Bedeutung*. But these remarks cannot, without qualification, be pressed to yield their natural implication. Since the use and understanding of nearly all expressions are vague and unperspicuous by his standards, the natural implication is that nearly all expressions lack a sense and denotation. But this is not the way Frege usually writes. In numerous places, he writes without qualification of the concept of number and of the sense and denotation of expressions in ordinary language and mathematics.

One could, of course, distinguish between a full-strength doctrine—that vague expressions do not fully express senses—and an approximate doctrine for public consumption—that expressions express senses (*sotto voce:* if only indefinitely or approximately). There is something to this point. But it still leaves Frege with the unexplicated notion of indefinite expression.

I think that Frege was guided by a more general conception, one that is substantially different from the ordinary notion of conventional meaning-expression. According to the ordinary notion of expression, vague terms do not "express" definite senses or denote definite concepts. But this is just a consequence of the point that vague understanding and usage do not fix or articulate definite senses or concepts. To say, as Frege says, that 'Number' *does* denote a concept and *does* express a sense is to say that the ultimate foundation and *justification* of mathematical practice supplements current usage and understanding of the term in such a way as to attach it to a concept and a sense. From this point of view, vague usage and understanding do not entail vague sense-expression.

Currently, we may have to guess at the sense (*PW* 13/*NS* 14, "Booles rechnende Logik und die Begriffsschrift" (1880–1881)). Even the best mathematicians may grasp it only haphazardly and intermittently (*PW* 222/*NS* 240, "Logik in der Mathematik" (1914)). Adequate, full grasp of what is "expressed" requires not just reflection on one's usage, but mathematical work and acquisition of mathematical knowledge (*PW* 33/*NS* 37, "Booles rechnende Logik und die Begriffsschrift" (1880–1881); *KS* 369, "Die Verneinung," O 150).

In spite of these peculiar points, there are two features of Frege's rationalism that allow him to count vague mathematical terms as "expressing" definite senses and denoting definite denotations. First, mathematical practice is held to be founded on a deeper rationale than anyone has previously understood. (This is, I think, the fundamental target of Wittgenstein's work on the philosophy of mathematics.) Second, Frege is attracted to construing the most fundamental sort of *a priori* knowledge on the model of insight or understanding. Thus the mathematical work of his logicist project is supposed to yield and be justified by the acquisition of insight—understanding of the true sense and denotation underlying ordinary mathematical uses of 'number' and appreciation of the self-evident, underlying principles of logic.

I cannot discuss the pros and cons of these features of Frege's rationalism here. There is no question that his deeply rationalist conception of sense-expression is highly idealized, and rather far removed from our ordinary notions of meaning, understanding, and thought. But the strong idealization is fully intended. As Descartes and Plato did before him, Frege emphasizes how difficult it is to attain full insight—full understanding (*FA* p. vii; *KS* 122, "Über das Trägheitsgesetz", O 158; *PW* 12–13, 222/*NS* 16–17, "Booles rechnende Logik und die Begriffsschrift" (1880–1881); *NS* 240, "Logik in der Mathematik" (1914)). Such understanding is not separable from the deepest sort of knowledge (*KS* 122–124; *PW* 33/*NS* 37, "Booles rechnende Logik und die Begriffsschrift" (1880–1881); *KS* 369, "Die Verneinung", O 150). We shall return to these points in Section VI.

The uncertainty expressed in passages (A) and (B) thus occurs within a larger rationale. Frege was not only uncertain or unclear about sharp boundaries for the pre-established usage of such terms as 'cardinal number'. He was also uncertain about the definition he gave for it in the *Foundations of Arithmetic*. Although a cursory reading of (A) and (B) suggests that Frege's uncertainty centered on matters of substance, it is a subtle question to what extent his worries were expositional and to what extent they were substantial.

II. PASSAGE (B)—CONTEXTUAL DEFINITION

Since, despite their kinship, passages (A) and (B) are indicative of very different tendencies in Frege's thinking, we shall discuss them separately. First, (B). In a letter to Russell, July 28, 1902, after discovery of the paradox, Frege writes that he had long struggled against recognizing courses of values, classes, or extensions of concepts, but that he has found no other answer to the question of how we grasp logical objects (numbers) than that we grasp them as extensions of concepts (*PMC* 140–141/*WB* 223).

There is evidence that Frege did struggle against using the notion of the extension of a concept in defining number. A manuscript from the period immediately after the publication of the *Foundations of Arithmetic* had the stated purpose of defining cardinal number without extensions of concepts, according to a list of Frege's *Nachlass* compiled by Heinrich Scholz.[6] Unfortunately, this manuscript was among those destroyed during an American bombing raid in 1943. Scholz's very brief summary of the manuscript makes it difficult to see how Frege reasoned. Besides the above-mentioned purpose

[6] Cf. Albert Veroort, "Geschichte des wissenschaftlichen Nachlasses Gottlob Freges und seiner Edition. Mit einem Katalog des ursprunglichen Bestands der nachgelassene-Schriften Freges" in *Studien zu Frege*, i, item 47, p. 95

of the manuscript, the summary merely mentions three parts of the manuscript with the following headings:

(a) Is it necessary to grasp numerical equality as strict identity?;
(b) Is it possible to define a judgeable content that contains $N_{\tau}F(\tau)$ [Frege's early expression for 'the number of F's'] in such a way that one says that the content may not change whenever F is replaced by G, as long as $F(\varrho)\underset{\varrho}{\smile}G(\varrho)$ [F and G are numerically equivalent];
(c) (The difficulty of the definition of an object through a recognition judgment [explicit definition]).

Scholz notes that (b) contains discussions of the definition of object and an early discussion of the definition of a concept's extension.

This is not much to go on. It appears that in (b) Frege was raising questions about the notion of object, as well as trying to define or explicate 'extension of a concept'. It also appears that he was considering a contextual definition of 'the number of F's' that would leave a containing sentence's "content" unchanged. Perhaps in (a) he was attempting to interpret the equals sign in arithmetic not as identity but as numerical equivalence. This latter notion he had defined (in *Foundations*) for the general case, in second-order logic. Perhaps in (c) Frege was reconsidering *Foundations* sections 56–61, 66–68, 107, in which he asserts that numerical expressions are singular terms and denies that a contextual definition of numerical expressions will yield a concept of number with sharp boundaries. His reason is that such definitions do not define whether an expression designates a number if it occurs outside the favored sentential contexts.

The difficulty for an explicit definition of numerical expressions, on the other hand, was that on Frege's view it forces one to take logic to be committed to the existence of particular objects, whose existence can be known *a priori* and on purely conceptual grounds. (cf. *FA*; secs. 3, 88–89, 104–105.) Given his exposure to Kantian themes, and his development of first- and second-order logic in 1879, Frege must have found this view unobvious, despite occasional, hopeful rhetoric to the contrary (*FA*, sec. 105). Moreover, his strictures on explicit definition required him to *demonstrate* the existence and uniqueness of any object for which a singular term was thus introduced (*FA* 106, 114–115; *BL* 11–12/*GG* i. 14). Even to someone unacquainted with the paradoxes and the incompleteness theorems, carrying out this task for extensions of concepts, for the general case, must have seemed a technically daunting task. (There is some suggestion of this in 'Über formale Theorien der Arithmetik' (1885), *KS* 110–111, O 103–104.) It was a task, of course, that Frege failed to fulfill (cf. *BL*, sec. 31; appendix II; *PMC* 132/*WB* 213, Russell to Frege, 6/16/1902). The proof for particular cases would depend on Frege's Basic Law (V), about which Frege expressed uncertainties before 1902, and which turned out to be inconsistent.

It seems reasonable then to conjecture that in the lost, post-*Foundations* manuscript Frege was reconsidering the whole question of whether numbers are objects—whether numerals and expressions like 'the number of F's' are primitive singular terms in arithmetic and in counting. He seems to have been contemplating a contextual definition of such singular terms (roughly in the spirit of Russell's "no class" theory). Perhaps he also considered an account of the equals sign as indicating not identity but numerical equivalence among concepts.

Responding to Russell's paradox in an appendix to the second volume of *Basic Laws*, 1902, Frege considers an alternative that bears some resemblance to the one he appears to have been considering fifteen years earlier. Frege considers treating terms for extensions (hence for numbers) as syncategorematic. Terms for extensions would be pseudo-singular terms, with no denotation. This is one way of construing the idea of providing singular numerical expressions with merely contextual definitions within logic. It may reflect his earlier investigation. Frege rejects this alternative on the ground that it would not allow one to explicate quantification into the position of such singular terms: "thereby, the generality of arithmetical propositions would be lost." He also claims that it would be incomprehensible on this view how one could speak of a number of classes (extensions) or a number of numbers (*BL* 129–130/*GG* ii. 255).[7]

Frege does not, in the appendix, consider recapturing the numbers as functions or concepts in the manner of second-order arithmetic, with no appeal to classes or extensions. It is clear, however, that this idea would run against Frege's grammatical grain. Frege seems to have remained committed to the view throughout the period that numbers were, if *anything*, the denotations of singular terms—objects. Concept expressions could never denote what singular expressions denoted (*FA*, sec. 51).

Scholz's notes on the lost manuscript are the only evidence I know of that Frege seriously considered a fundamentally different way of defining singular numerical expressions. It is not very strong evidence. Still, in my view, it is probable that he was considering the possibility of contextual definition even in writing *Foundations* and that this is the sort of alternative to "bringing in extensions of concepts" that he had in mind in passage (B) (*FA* 117). For in

[7] In response to a suggestion of Russell in correspondence, he also considers taking extensions to be "improper objects," "not objects in the full sense," for which "the law of excluded middle fails to hold." The idea is to restrict the instantiation from quantified sentences to singular terms denoting extensions in accordance with rules governing levels, in the spirit of Russellian type theory (*BL* 128–129/*GG* ii. 254–255. Cf. also letter to Russell, 9/23/1902; *PMC* 145 ff.). Frege rejects this alternative because of its complication and unintuitiveness. He particularly objects to stratifying identity. I think it unlikely that he would have considered this alternative seriously prior to discovering the paradoxes. A sort of type theory was proposed by Schröder before the discovery of the paradoxes. But this flowed from a conception of classes that was quite different from Frege's. Cf. Alonzo Church, "Schröder's Anticipation of the Theory of Types" *Erkenntnis*, 10 (1976), 407–11. Frege strongly rejected both Schröder's and Russell's type theories (*KS* 197–198, "Kritische Beleuchtung einiger Punkte in E. Schröders Vorlesungen über die Algebra der Logik", O 439–441; *BL* appendix II).

that passage he is specifically concerned with the difficulties of explicit definition.

III. PASSAGE (A)—CONCEPT AND OBJECT

We have more evidence for interpreting passage (A). Indeed, I think it possible to say with some confidence what is meant. Frege referred to the footnote on three other occasions. We shall discuss two of these in this section, saving the third for section IV. By scrutinizing these retrospective remarks, one can see that in (A) Frege was alluding not to a fundamentally different method of defining numerical expressions, but to uncertainty over using the term "extension of a concept," as opposed to another expression with substantially the same ontological implications.

(A) is discussed by Frege in "On Concept and Object" (1892). There he writes:

(D) If we keep in mind that, in my way of speaking, expressions like "the concept F" signify not concepts but objects, Kerry's objections will already for the most part collapse. If he thinks . . . that I have identified concept and extension of a concept, he errs. I merely expressed my view [presumably in (A)] that in the expression "the Number that attaches to the concept F is the extension of the concept numerically equivalent with the concept F" one could replace the words "extension of the concept" by "concept." Notice carefully here that this word [the word "concept"] is combined with the definite article. Besides, this was only an incidental remark on which I based nothing (G & B 48/*KS* 172, O 198–199).

This passage immediately suggests how in 1884 Frege intended to answer the first of the two objections he raises in (A) to the substitution: the substitution does not contradict his earlier statement that numbers are objects because "the concept *numerically equivalent with the concept F*" denotes an object not a concept.[8]

Which object is, of course, not immediately clear. But there is reason to think that Frege was not here contemplating serious ontological divergence from the definition in terms of extensions of concepts. The substitution is justified by an identity. In a draft of "On Concept and Object" published in the *Nachlass* (written between 1887 and 1891), in the course of a passage overlapping the end of (D), Frege comments again on (A):

(E) Besides, this was only an incidental remark on which I based nothing, in order not to have to grapple with the doubts to which it could give rise. So Kerry's opposition to it does not at all bear on the core of my position.* (*Whether

[8] Benacerraf, "Frege: The Last Logicist", 30 and Terence Ward Bynum, "The Evolution of Frege's Logicism", in Schirn (ed.), *Studien zu Frege*, 282, infer from (A) that Frege was contemplating replacing extensions of concepts with concepts.

> instead of the expression "the extension of the concept" one should simply say "the concept," I see as a question of expediency.) (*PW* 106/*NS* 116, "Eine kritische Auseinandersetzung mit Kerry" (1891–1892))

The remark about expediency (couched, as it is, in the formal mode) suggests that the issue on Frege's mind in passage (A) was partly expositional—how best to present the definition of number that he proposed in *Foundations*, section 68. The remark also suggests what object Frege supposed the concept *horse* to be. It was supposed to be the extension of the concept *horse*.

Although in 1884 Frege probably thought that 'the concept F' and 'the extension of the concept F' denoted the same object, he probably had not decided what extensions of concepts (or even concepts) were. Not until later (1891 in "Function and Concept") did he introduce the notion of courses of values, with which he identified extensions of concepts (cf. *BL* 5–6/*GG* i. 9–10). The expositional uncertainty expressed in (A) is not entirely separable from this ontological uncertainty.

IV. PASSAGE (A)—ONTOLOGICAL OPTIONS

To understand Frege's intellectual situation in 1884, one must remember that he was faced with a wide variety of uses, among respected authors, of such terms as 'extension', 'concept', 'class', 'set', 'system'. Not only usage, but conception varied. The variations were symptomatic of lack of clarity. Explications of the notions of set, even by such men as Cantor and Dedekind, now strike one as remarkably vague and unstandardized. It was a time of conceptual ferment.

More specifically, Frege was confronted, on one hand, by an ancient logical tradition of using 'extension' in discussions of natural language, and, on the other, by a newer tradition among mathematicians (represented in Frege's thinking chiefly by Cantor, Schröder, and Dedekind) that was attempting to define numbers in terms of sets (classes, systems). I will discuss these two traditions briefly.

It is likely that Frege was influenced in his use of the term 'extension' ('*Umfang*') by the Boolean logical tradition which, in turn, derives from Leibniz. Frege associates the term with Boole in unpublished work dating from 1880–1881 (*PW* 15–16, 33/*NS* 16–17, 37, "Booles rechnende Logik und die Begriffsschrift" (1880–1881); cf. *LW* 52). The term is also used in Mill's *System of Logic* (*SL* 3–4) and Jevons's *The Principles of Science* (*PS* ii. 2)—both of which Frege cites in *Foundations*. Extension of a concept (*Begriffsumfang*) is also used by Schröder—interchangeably with 'class' (*Klasse*) (e.g. *AL* i. 222, 233—Schröder gives space to expounding Mill). Jevons, for example, writes "...every general name causes us to think of some one or more objects belonging to a class. A name is said to denote the object of thought to which it may be applied . . . the objects denoted form the extent of meaning of the term; the qualities form the intent of meaning" (*PS* ii. 2). Jevons's statement brings out the emphasis, within this logical tradition, on a dual meaning for concept words. Each sort of meaning (property and extension) was indicated, in some sense, in every use of a general term.[9]

The tradition tends to emphasize that the intension of a concept or term is "more basic" than its extension since the qualities, marks, or characteristics that form the intension define or fix the extension (and not vice-versa) and because the extension is thought only "through" such characteristics (cf. *LL* 102; *SL* v. 4; *PS* xxxi. 14). It is also sometimes argued, apparently in view of context-dependent constructions, that the intension is more basic because extension shifts though the intension remains constant (cf. *SL* v. 3; *PS* iii. 7). This is, of course, an argument that left Frege cold—perhaps because it is inconsistent with the view that intensions fix extensions. But the semantical and epistemic arguments for the priority of intensions must have struck a responsive chord.

This view of the semantical and epistemic priority of intensions—properties, qualities, characteristic marks—over classes was opposed, at least implicitly, by the development of set-theoretic approaches to number theory by authors like Cantor and Dedekind. There was a fair amount of contempt among mathematicians for traditional logic—particularly, for its lack of precision and its fruitlessness. Frege refers to this attitude occasionally (*PW* 34/*NS* 38, "Booles rechnende Logik und die Begriffsschrift" (1880–1881); *KS* 104, "Über formale Theorien der Arithmetik", O 95–96). And Cantor exhibits it in his 1885 review of *The Foundations of Arithmetic* (*CGA* 400), with a disparaging reference to 'extension' and the old "school logic."

The beginnings of an analysis of number within intuitive set theory had been carried out at the time Frege wrote *Foundations*. Frege must have learned from this technical work. In fact, in 1885 Frege credits Cantor's 1885 review (and less explicitly in *FA* 97–98, Cantor's 1883 paper) with an equivalent definition of 'the number of F's' (*KS* 112, "Erwiderung auf Cantors Rezension der *Grundlagen der Arithmetik*"; cf. *CGA* 441, 167; *FA* 97–98).[10]

[9] George Boole, *Collected Logical Works (LW)*, ii. (Chicago: Open Court Publishing Company, 1940), (*The Laws of Thought* was published in 1854); John Stuart Mill, *A System of Logic* (*SL*) (London: Longmans, Green and Co., 1865); W. Stanley Jevons, *The Principles of Science* (*PS*) (London, Macmillan and Co., 1920)—first published, 1865; Ernst Schröder, *Vorlesungen über die Algebra der Logik* (*AL*), Bd. I (Bronx, NY: Chelsea Publishing Company, 1966). The term 'extension' has a long history deriving from the Latin '*extensive*'. Sir William Hamilton, *Lectures on Logic*, ii (New York: Sheldon and Company, 1858), 100–1, traces the term through the *Port Royal Logic* (1718) to Cajetan (1496) back to Boethius. He thinks it has roots in Aristotle.

[10] Cantor's definition, in his 1885 review of Frege's *Foundations*, goes back to Cantor's *Grundlagen einer allgemeinen Mannigfaltigkeitslehre*, 1883. Cf. Georg Cantor, *Gesammelte Abhandlungen* (*CGA*), ed., E. Zermelo, (Hildesheim: Georg Olms, 1962). In a note on the review, *CGA* 441–2, Zermelo says that the definitions are equivalent. Counting them so involves ignoring the deep differences between Cantor's and Frege's notion of class. Zermelo is aware of this.

Despite his likely respect for the mathematical work in this tradition, Frege was, from very early on, dissatisfied with its conceptual foundations. In particular, he criticized that notion of class or set that derives from grouping together elements—the ancestor of the iterative conception of set. He associates the notion with a "childish" Millian viewpoint, according to which the elements of a set are seen simply as physical parts of a physical whole (*KS* 105, "Über formale Theorien der Arithmetik", O 96; *FA* secs. 22–25, 28).[11] He points out that any conception of class or set that would be relevant to explicating what numbers are would have to include events, methods, and concepts. And he holds that these are not reasonably seen on the part–whole model (ibid.). He further believes that attempts to attain an appropriate level of abstraction, sufficient to deal with the wide applicability of the notion of number typically lead to psychologistic accounts of grouping the elements: accounts that appeal to the mind's abstracting from physical properties and the like (*FA*, secs. 25–28, 50–52; *KS* 164, "Rezension von: Georg Cantor, Zur Lehre vom Transfiniten", O 269; *PW* 69–71/*NS* 77–80, "Entwurf zu einer Besprechung von Cantors Gesammelten Abhandlungen zur Lehre vom Transfiniten" (1890–1892); *BL* 29–31/*GG* I. 1–3). After 1884 he particularly indicts Dedekind and Cantor on this charge.[12]

Three of Frege's criticisms of the "element-grouping" conception of set are independent of his association of it with Mill's "agglomerative" conception and with psychologistic accretions. Frege repeatedly appeals to the criticisms in the period 1880–1895 (as well as later). The primary objection derives from the tradition in logic, mentioned earlier, according to which intensions (properties, marks, characteristics) are prior to extensions. More specifically, for Frege, predication is epistemically and logically prior to abstract objects. Frege's point is that the elements of a class are fixed, "delimited," only through "concepts." Elements are what fall under the concept. The relevant elements are determined, held together as a totality, only through "characteristics" they have in common—through a rule governing elementhood (*PW* 34/*NS* 38, "Booles rechnende Logik und die Begriffsschrift" (1880–1881); *FA* 59–67; *KS* 105, "Über formale Theorien der Arithmetik", O 96–97; *KS* 164, "Rezension von: Georg Cantor, Zur Lehre vom Transfiniten", O 270; *KS* 210, "Kritische Beleuchtung einiger Punkte in E. Schröders Vorlesungen über die Algebra der Logik", O 455–456; *BL* 30/*GG* i. 2–3; G & B 105/*KS* 209, "Kritische Beleuchtung einiger Punkte in E. Schröders Vorlesungen über die Algebra der Logik", O 454–455). Frege

[11] For a nonhistorical discussion of this notion, see my "A Theory of Aggregates", *Noûs*, 11 (1977), 97–117.

[12] There is good reason to think that Frege was right on the *ad hominem* point. Cantor's informal remarks about sets carry a remarkably strong, though probably not entirely intended, suggestion of dependence of sets on human minds (*CGA* 283, 387–8, 411, 413 n, 440). Similar points apply to Dedekind's remarks about sets and numbers—except that Dedekind's psychologism seems more deeply meant. (Cf. Richard Dedekind, *Gesammelte Mathematische Werke*, Bd. III, ed. R. Fricke, E. Noether, and Ö. Ore (Braunschweig: Freidr. Vieweg & Sohn Akt.-Ges., 1932), 335, 344 ff.)

explicitly associates concepts with properties and characteristic marks—both terms from the Boolean–Millian logical tradition.

I think that behind the objection lies Frege's assumption of a context principle according to which the justification for postulating abstract entities, such as sets, extensions, numbers, concepts and the like, derives from their role in providing denotations or meanings (*Bedeutungen*—in the earlier writings, contents, *Inhalte*) *for components* of *sentences* that express true thoughts. Frege held that one could think about such objects only through predication or nominalizations of predication; one could not justify the postulation of abstract entities by making lists. Such a procedure, he writes, is "very arbitrary and in actual thinking without significance" (*PW* 34/*NS* 38, "Booles rechnende Logik und die Begriffsschrift" (1880–1881)). An evaluation of the context principle and Frege's objection is, however, beyond the bounds of this paper.

Two other objections are in effect corollaries of the first. One is that the element-grouping conception cannot account for the null class. By contrast, the idea that there are concepts under which no entities fall is familiar and natural (*PW* 34/*NS* 38 "Booles rechnende Logik" (1800–1881); *FA* 30, 59, 64; *KS* 105, "Über formale Theorien der Arithmetik", O 97; G & B 102/*KS* 206–207, "Kritische Beleuchtung einiger Punkte in E. Schröders Vorlesungen über die Algebra der Logik", O 451). The other objection is that the notion of an infinite, completed totality, even a totality with the cardinality of the first number class, cannot be justified by "logical addition." Human beings are inevitably finite, and "inner intuition" is too subjectivistic and unclear a notion to use in justifying such totalities. "Going on in the same way," he might have added, is too indefinite a notion to explicate the objectivity and seeming mathematical definiteness of these large totalities. Such totalities could be justifiably seen only as deriving from definite concepts (rules of application) indicated in thought (*PW* 34/*NS* 38 "Booles rechnende Logik" (1880–1881); *KS* 164, "Rezension von: Georg Cantor, Zur Lehre vom Transfiniten", O 270; *BL* 31; *GG* i. 2–3; cf. also *FA* 98).

Frege was guided by these objections in insisting that any abstract, class-like objects in terms of which the numbers were to be defined had to be derivative from concepts, which he had regarded since 1879 not only as properties or characteristic marks, but also as a special sort of function (G & B 12–14/*B* 15–17).[13] So extensions of concepts would have to be, in some sense, objects derivative from concepts. Yet they could not *be* concepts because of Frege's insistence that an object cannot be predicative: concepts were associated only with predicates (or other function expressions); objects were denoted only by singular terms. I shall not attempt to explain this notorious insistence, though we shall return to it (cf. G & B 42–55/*KS* 167–178, "Über Begriff und Gegenstand", O 192–205).

[13] Hans Sluga, *Gottlob Frege* (London: Routledge and Kegan Paul, 1980), 81, emphasizes the point and does much else to show how many of Frege's deepest conceptions derive from the 1879 *Begriffschrift*.

Bearing in mind Frege's attitudes to the logical and mathematical traditions, let us return to the question of why in (A) Frege considered using 'the concept *F*' as an alternative to 'the extension of the concept *F*', to denote the same thing.

One reason may have been that Frege saw 'the concept *F*' as the natural expression, within his logical theory, for the notion traditionally but confusingly expressed by "the concept in extension." Extensions were objects—the referents of singular terms—but they were also spoken of in the logical tradition as aspects or guises of concepts. Another reason might have been that the substitution would have been more palatable to a mathematical tradition suspicious of the term 'extension', but used to talking of functions in an extensional way (i.e. in such a way that functions were counted the same if for the same arguments they had the same values). A trace of this motivation may be seen in a long footnote to a discussion of Basic Law V in the second volume of *Basic Laws*:

> Few mathematicians will reflect upon expressing the circumstance that f() always has the same value as g() for the same argument as "f = g". Nevertheless, the error involved in this springs from an inadequate conception of the essence of the function. . . . Although accordingly the designation "f = g" cannot be recognized as correct, it shows that the mathematicians have already made use of the possibility of our transformation [in Basic Law V]. (*GG* ii. sec. 147)

Given Frege's long-standing inclination toward identifying concepts with certain functions, the substitution may have seemed expositionally attractive. The expression 'the concept . . . ' itself was sometimes used by Cantor, and no doubt others to refer to abstractions.[14] Frege may have seen in the expression a means of enforcing his insistence that classes are derivative from (unsaturated) concepts.

In fact, Cantor uses the expression in almost precisely the same context in which Frege used 'the extension of the concept . . . ' in defining 'the number of F's' in *Foundations*, 79–80. Frege's third reference to passage (A) occurs in a reply to Cantor's uncomprehending 1885 review of *Foundations*. Cantor had criticized Frege for defining 'the number of F's' as the extension of *F* (*CGA* 440). Frege (1885) points out that this was not his definition. He makes reference to Cantor's definition of 'power' (*Mächtigkeit*), which for finite cardinal numbers is equivalent to Frege's definition of 'Number' ('*Anzahl*') (if one ignores Frege's special doctrines about what concepts are—cf. note 10 above). In effect, Frege alludes to the equivalence and adds:

[14] Charles Parsons pointed out in conversation that Cantor consistently uses 'Begriff' for the cardinality of a set and 'Menge' or other terms for sets. One might speculate that Cantor had at least some sense of the modern view which demands provision for a distinction of levels. Frege's distinction between class and concept and his assimilation of classes to extensions of concepts seem to flow from an entirely different motivation.

(F) Incidentally, the difference that Mr. Cantor writes 'General Concept' ('*Allgemeinbegriff*') where I write 'extension of the concept', seems inessential as noted in the footnote, p. 80 [i.e. passage (A)] in my work. (*KS* 112, "Erwiderung auf Cantors Rezension der *Grundlagen der Arithmetik*").

Cantor's definition derives from a paper of 1883 that Frege cites in *Foundations*. It is thus possible that Frege's mention of the substitutability of 'the concept' for 'the extension of the concept' in *Foundations*, passage (A), was a not very explicit acknowledgement of the equivalence of Cantor's 1883 definition (cf. n. 10; also *Foundations*, 97–98). The definition of number in terms of numerical equivalence was a key and original feature of Cantor's program, and it is possible that Frege got the idea from Cantor. Later, in 1892, Frege decided that Cantor's notion of class was too different from his own to count the definitions equivalent (*KS* 164, "Rezension von: Georg Cantor, Zur Lehre vom Transfiniten", O 270).

The interchange with Cantor and the fact that in 1884 Frege already firmly distinguished concept and object suggest how Frege might have answered the second of the two objections he raises in (A) against his view that 'the concept' could replace 'the extension of the concept'. The objection was that concepts can have the same extension without coinciding. Two complementary answers were available to Frege. First, as Furth points out, Frege used 'coincide' ('*zusammenfallen*') equivalently with 'are identical' (*BL* pp. xiii–xiv; *KS* 184, "Rezension von: E. G. Husserl, *Philosophie der Arithmetik*, i", O 320). Since identity is a relation between objects, not concepts, concepts can indeed have the same extensions without themselves being identical. But then since 'the concept ...' denotes an object—a class or extension, not a concept (*FA* 63)—the objection is irrelevant. Second, Frege may, by 1884, have been contemplating a mathematician's extensional view of concepts and functions: no discrimination of concepts without discrimination among objects that fall under them. Thus concepts with the same objects falling under them stand in the analog for concepts of the relation of identity. Concept expressions with the same objects falling under them are intersubstitutable *salve veritate*. At any rate, this is an answer he would have given by 1891.[15]

[15] There is some evidence that before *Foundations* Frege had adopted from the tradition in logic a more intensional view of "concepts": (*PW* 18/*NS* 19–20, "Booles rechnende Logik und die Begriffsschrift" (1880–1881)). By 1891 at latest, he was propounding the extensional view, sometimes expressly noting its congeniality with the writings of "extensionalist logicians" (G & B 21/*KS* 125, "Funktion und Begriff", O 1; G & B 80/*KS* 183–184, "Rezension von: E. G. Husserl, *Philosophie der Arithmetik*, i"; O 319–320; *PW* 118 ff./*NS* 128 ff, "Ausführungen über Sinn und Bedeutung" (1892–1895); *BL*/*GG*, secs. 1–4, 9–10, 36; *GG* ii. 146–147). The evidence regarding Frege's position on the matter in *Foundations* is sketchy. But a number of considerations suggest that under pressure from mathematical considerations, Frege might have already found the extensionalist conception attractive. (Cf. (A), the exchange with Cantor (E), secs. 63–65, and *PW* 16/*NS* 17, "Booles rechnende Logik und die Begriffschrift" (1880–1881).)

The development of the *Sinn–Bedeutung* distinction after 1884 certainly made it easier for Frege to adopt an extensional view of concepts. Or perhaps the latter helped motivate the former. The *Sinn* of a

Let us summarize the view of passage (A) that has so far emerged. Frege appears to have regarded the substitution of 'the concept' for 'the extension of the concept' as ontologically insignificant. Both denoted objects. The terminological ambivalence reflects an uncertainty about how to present his logicist theory, but it also reflects unclarity about what numbers are. This unclarity stemmed partly from unclarity about what concepts are. Whereas it is debatable whether in 1884 Frege already viewed concepts extensionally, it is virtually certain that he did not yet view them as functions whose values are truth-values (*BL* 5–6/*GG* i. 9–10). Without this view, he could not yet have developed his explication of what extensions of concepts are.

Frege was dissatisfied with the "element-grouping" conception of classes, but had not yet developed a satisfying alternative. In this context, he was exposed to Cantor's 1883 definition of 'power' (roughly, 'cardinal number'). It is unclear whether Frege developed his own definition independently, or whether his definition made use of Cantor's. At any rate, in writing *Foundations* he was confronted with a technically similar definition couched in a somewhat different vocabulary—a vocabulary that highlighted difficult issues about the relation between concept and object, and the very meaning of 'concept' and 'extension of a concept'. In passage (A), Frege notes the alternative formulation and indicates that he thinks the difference is not substantial. He contemplates maintaining the strict distinction between concept and objects, while claiming that 'the concept F' denotes an object—presumably the same object as does 'the extension of the concept F'. What this object is was to be decided by further work. As passage (B) suggests, Frege was not even sure that he wanted to stand by his appeal to objects at all, much less what objects "extensions of concepts" were to be.

V. JUSTIFYING LAW (V)—RETREAT FROM PASSAGE (A)

Let us turn to the role in all this of Basic Law (V), the axiom governing extensions of concepts that led to contradiction. Frege's willingness to substitute 'the concept F' for 'the extension of the concept F' in passage (A) reflects more than an implicit acknowledgment of Cantor's different terminology in defining 'Number'. It also reflects Frege's struggle to justify Law (V) as a logical law. This struggle was inextricably bound up with the attempt to arrive at a satisfactory conception of logical objects—extensions of concepts.

concept word could fulfill the cognitive role that motivated intensionalist views of concepts, and *Bedeutung* could fulfill the mathematically and semantically fundamental role of contributing to truth-value. It would be interesting to know the exact developmental story about the relation between Frege's views on concepts and on *Sinn–Bedeutung*. (Cf. *PW* 112–123/*NS* 133–134, "Ausführungen über Sinn und Bedeutung" (1892–1895).)

The discussion of equivalence class definitions and the analogy between direction and number in *Foundations* (*FA*, secs. 62–67) suggest that in 1884 Frege may have already had in mind the form of Law (V). In fact, he states the basic idea as early as 1880–1881, attributing it to the Boolean tradition (*PW* 16/*NS* 17, "Booles rechnende Logik und die Begriffsschrift"). The law can be formulated as follows:

(V) $(x)(Ax \leftrightarrow Bx) \leftrightarrow \grave{\epsilon}A(\epsilon) = \grave{\alpha}B(\alpha)$

(where '$\grave{\epsilon}A(\epsilon)$' may be read, 'the extension of the concept A'). In 'On Function and Concept' (1891), Frege claims that the two sides of an equivalence that is formally similar to (V) have the same *sense* (G & B 26–7/KS 130–1, "Funktion und Begriff", O 10–11): '$(x)(x^2 - 4x = x \cdot (x - 4))$' is said to have the same sense (*Sinn*) as '$\grave{\epsilon}(\epsilon^2 - 4\epsilon) = \grave{\alpha}(\alpha \cdot (\alpha - 4))$'. In the same article Frege assimilates concepts to functions. So there is no deep difference between this remark about functions and an analogous remark about concepts. Moreover, this is the article in which Frege first introduces the *Sinn–Bedeutung* distinction. So it appears that Frege was conceiving at least instances of (V), and probably (V) itself, as expressing *the same sense* on the two sides of the biconditional. Such a conception, if correct, would certainly justify (V) as a logical law. For the biconditional would have the same sense and same self-evidence as '$(x)(Ax \leftrightarrow Bx) \leftrightarrow (x)(Ax \leftrightarrow Bx)$', and would not depend for its truth on any notions or truths in a special science (cf. *FA* 4; *BL* 127/*GG* ii. 253).

What led Frege to this implausible view? There seem to be two possible, mutually compatible, sources. One is the logical tradition, which we discussed earlier, in which each sentence containing a general term, and each general term, had a dual "meaning"—an "intensional" and an extensional one. Frege was clearly influenced by this tradition. He took his term 'extension of a concept' from it. Very probably his insistence on the priority of concepts (which he originally viewed as properties, probably interpreted intensionally) over classes was influenced by it. As he came to view concepts extensionally and developed the *Sinn–Bedeutung* distinction, he may have retained the view that (even in *addition* to a sense) the use of concept words fixed a dual "meaning"—concept and extension of the concept (cf. *PW* 122–123/*NS* 133–134, "Ausführungen über Sinn und Bedeutung" (1892–1895)). Although they "meant" (*bedeuteten*) only concepts, their use fixed an extension as well. In natural language, a given sentence with a given sense determined both *Bedeutungen* (cf. G & B 49–50/*KS* 173–174, "Über Begriff und Gegenstand", O 200–201). In Frege's formal notation, distinct sentences would be used in denoting the distinct *Bedeutungen*, but they would express the same sense—only "in different ways" (G & B 27, *KS* 130, "Funktion und Begriff", O 10–11; cf. *FA*, sec. 64).

The second possible source of motivation for seeing the two sides of Law (V) as expressing the same sense derives from Frege's use of 'the concept F'.

As far back as 1879, Frege had fallen into the rather natural habit of reading sentences of the form F(*a*) as '*a* falls under the concept *F*' or G(*a, b*) as '*a* stands in the *G* relation to *b*' (*B*, sec. 10; *PMC* 101/*WB* 164–165, Frege to Marty, 8/29/1882; *FA*, secs. 55, 58, 74–75; G & B 30, 46–47/*KS* 133, "Funktion und Begriff", O 15; *KS* 171, "Über Begriff und Gegenstand", O 197; *BL*/*GG*, sec. 4). This habit continued through the publication of *Basic Laws*. It is clear that he takes the two modes of expression to have the same sense. Perhaps at first Frege regarded 'the concept F' (and 'falls under' in 'falls under the concept F') as syncategorematic. But passage (A) suggests that in 1884 he took 'the concept F' to be referential. This view is quite explicit in "On Concept and Object" (1892). Frege writes that 'the concept *man*' denotes an object and that in 'Jesus falls under the concept *man*', 'the concept *man*' is only part of the predicate. The whole expression 'falls under the concept *man*' means (*bedeutet*) the same as 'a man' (G & B 46–47/*KS* 171, O 197; cf. also G & B 45/*KS* 170, O 196). Similarly, at the end of the essay, Frege writes that in the sentence 'the number 2 falls under the concept *prime number*', 'falls under' expresses a sense that is two ways unsaturated and denotes (*bedeutet*) a relation (between objects) (G & B 54–55/*KS* 178, O 205).

It is immediately after the discussion of 'Jesus falls under the concept *man*' that Frege refers back to passage (A) in passage (D). He then claims that

(i) There is a square root of 4

and

(ii) The concept square root of 4 is realized

express the same thought. But he holds that the first six words of (ii) denote an object, whereas 'is a square root of 4' in (i) denotes a concept. Similarly, the quantifier in (i) does not denote a concept co-extensive with what is denoted by 'realized' in (ii). (The former concept is second-level; the latter, first-level.) Frege then insists that a thought can be split up in many ways and that the same *sentence* (*Satz*) (as well as the thought it expresses) can be regarded as an assertion about a concept (*square root of 4*) and also an assertion about an object (the concept square root of 4)—"only we must note that what is asserted [assertively predicated] is different" (G & B 49/*KS* 173, "Über Begriff und Gegenstand", O 200; cf. also G & B 50/*KS* 173–174, O 200–201). This remark sheds interesting light on Frege's conception of the context principle (or syndrome of principles) at this stage in his career. It is also reminiscent of the "dual meaning" aspect of extension–intension theory discussed earlier.[16]

[16] There is a noteworthy parallel between this doctrine of 'falls under' and Frege's doctrine of truth. In this same year (1892, 'Über Sinn und Bedeutung', G & B 63–64/*KS* 149–150, O 34–35), Frege holds that an attribution of truth to a thought does not express a different sense from the thought itself: the relation between a thought and truth is that of sense to *Bedeutung*. Thus (iii) p, and (iv) p = the

In '2 falls under the concept *prime number*', 'falls under the concept *prime number*' denotes the same concept on Frege's theory of 1892 as does "$\cap\grave{\varepsilon}$Prime Number(ε)," since the expressions are co-extensive. This latter expression (defined in *BL* 34) could be roughly read 'falls in the extension of the concept *prime number*'. Here it would seem that 'falls in' is co-extensive with 'falls under' and 'the extension of the concept *prime number*' denotes the same object as 'the concept *prime number*'.

So it appears that in 1891–1892, Frege was still working out the implications of the substitution of 'the concept . . . ' for 'the extension of the concept . . . ' contemplated in passage (A). I conjecture that he was at least considering the possibility that the ordinary language paraphrase of F*a* into "*a* falls under the concept *F*" would lend support to the view that Basic Law (V) expresses the same sense on each side of the biconditional. Expressions of the form 'the concept *F*' would denote the extensions or courses of values that Frege needed to define the numbers.

Sometime soon after the publication of "On Concept and Object" and "On Sense and Denotation," Frege rejected this line of thought. In "Comments on Sense and Denotation" (dated in the period 1892–1895), Frege recommends that expressions of the form 'the concept F' be rejected (*PW* 122/*NS* 132–133, "Ausführungen über Sinn und Bedeutung") because "the definite article points to an object and belies the predicative nature of the concept." He goes on to state Law (V) in terms of extensions. He does not, however, deny that such expressions as 'the concept F' denote objects.[17] What he

True, express the same thought, even though an identity relation is denoted in (iv) but *seems* not to be in (iii). This can be compared to the doctrine of 'On Concept and Object' that (v) Fa, and (vi) *a* falls under the concept *F*, express the same thought even though a two-place relation is denoted in (vi) but seems not to be in (v). Apparently, according to the views of 'On Concept and Object', the same objects and concepts are, if not denoted, at least "fixed" in both (v) and (vi). Frege is not explicit about whether the point applies to (iii) and (iv), but I see no reason to think that he treated the cases differently.

This doctrine as applied to (v) and (vi) easily leads, by reasoning similar to that in G & B 54–55/*KS* 177–178, "Über Begriff und Gegenstand", O 204–205, to the conclusion that an infinity of objects and concepts are "fixed" within any given sentence and any given thought. For it appears that on the doctrine, (vi) has the same sense as

(vii) *a* stands in the relation of falling under the concept *F* to *b*.

Here '—— stands in——to——' is a three-place relation among three objects. Unless some means of restricting new concepts can be produced, an infinity is generated. "Dual meaning" would be an understatement.

[17] In fact, in a footnote on the same page, he says that they do: "These objects have the names 'the concept ϕ' and 'the concept X'." Earlier, however, in the same essay he may be expressing doubt: "the denotation of the expression 'the concept *equilateral triangle*' (insofar as one exists) is an object" (*PW* 119–120/*NS* 130, "Ausführungen über Sinn und Bedeutung" (1892–1895)). Sluga, *Gottlob Frege*, 142–3 claims that in this essay Frege "no longer holds ['the concept F'] to refer to anything." This claim is not supported by the evidence. Sluga also reads (A) as indicating that even when he thought the expression denoted an object, Frege never took the object to be the extension of the concept F. Sluga provides no support for this view. It is true that Frege never explicitly proposes the identification. But our evidence suggests that he made use of it, at least tentatively.

rejects is use of the expressions *in logic*. The reason he gives was fully available to him when he wrote "On Concept and Object." But there is no evidence until the present essay that he was willing to forego such use in logic. It is noteworthy that Frege's reason may be interpreted as centering on expositional considerations. The potential for misunderstanding that Frege had anticipated in (A) was realized in Kerry's objections. In any case, Frege gave up whatever hopes he may have had for using his views about 'the concept F' to strengthen the intuitive justification of Basic Law (V).

Although Frege uses 'the concept F' in an informal expositional way in *Basic Laws* (e.g. *BL*, sec. 3), there is no suggestion that the expression could be put to serious scientific use. Frege warns that it is misleading; he says that it does not denote a concept; and he omits to say that it denotes an object (*BL*, sec. 4, footnote). It is quite possible that by 1893 he regarded the expression as syncategorematic.

As far as I can find, however, he does not explicitly assert this view until several years after the discovery of the paradox. In a letter to Bertrand Russell of July 28, 1902 (*PMC* 141/*WB* 224), he says that the expression 'is a concept' is "logically speaking really to be rejected." He notes that it should name a second-level concept [at least!], but it presents itself "linguistically" as a name of a first-level concept. But even here, Frege does not explicitly say that 'the concept *F*' fails to denote. In 1906, he does so.

> Language stamps a concept as an object, since it can fit the designation of a concept in its grammatical framework only as a proper name. But it thereby really commits a falsification. Thus, the word 'concept' is itself strictly speaking already defective, in that the words 'is a concept' demand a proper name as grammatical subject; for they thereby demand a contradiction, since no proper name can designate a concept; or perhaps better still a piece of nonsense (*Unsinn*).... In the sentence 'Two is a *prime number*', we find a relation designated, that of subsumption. We can also say, the object falls under the concept *prime number*.... This makes it seem as if the relation of subsumption were a third element supervenient on the object and the concept. This is not the case; rather the unsaturatedness of the concept brings it about that the object, in effecting the saturation, sticks immediately to the concept, without needing any special binding agent. (*PW* 177–178/*NS* 192–193, "Über Schoenflies: die logischen Paradoxien der Mengenlehre" (1906); cf. also *PW* 193/*NS* 210, "Einleitung in die Logik" (1906))

Here, Frege seems to concede, in effect, that the sentence 'The concept *horse* is not a concept' is not true but senseless. He also counts 'falls under' syncategorematic, and thereby does the same for 'the concept F' in those contexts where it has any use at all. From 1914 on, Frege makes similar even more explicit claims that 'the concept *F*' fails to denote anything (*PW* 238–239, 249–250, 255, 272/*NS* 257–258, 269, "Logik in der Mathematik" (1914); *NS* 275, "Aufzeichnungen für Ludwig Darmstaedter" (1919); 291, "Erkenntnisquellen der Mathematik und der mathematischen Naturwissenschaften" (1924–1925)).

Thus it is possible that Frege held that 'the concept *F*' does not denote anything (as distinguished from merely being too misleading to include in scientific contexts) only after Russell's paradox showed that 'the extension of the concept *F*' is itself defective. Whether or not this is true, 'the concept *F*' is not said to be substitutable for 'the extension of the concept F' in *Basic Laws*. Passage (A) is, in effect, retracted.

Frege nowhere claims in *Basic Laws* that the two sides of Law (V) express the same sense. This omission may not be significant. He does claim that (V) is a logical law (*BL* 4/*GG* i. 7) and that it is "what one thinks when one speaks of extensions of concepts."[18] The omission may simply reflect a consciousness that the proposed law was potentially controversial, in that it was admittedly less obvious than the others (*BL* 3/*GG* i. 7; *BL* 127/*GG* ii. 253). In view of Kerry's objections, claims about sameness of sense may have seemed to Frege an unnecessary incitement to doubt. Still, it is striking that in *Basic Laws*, Frege foregoes any serious justification of Law (V), and even expresses doubts about it.

VI. LAW (V)—INCOMPLETE UNDERSTANDING

There is no question that Frege harbored uncertainties about Law (V) in 1893. They surface twice in the opening pages of *Basic Laws* (*BL* 3–4, 25/*GG* i. pp. vii, xxvi). And in the appendix, reacting to the paradox, he writes, "I have never concealed from myself its lack of self-evidence which the other [Laws] possess, and which must properly be demanded of a law of logic" (*BL* 127/*GG* ii. 253).

Why was Frege uncertain? It is clear that he was primarily concerned about extracting objects (the extensions) from concepts (*BL* 3–4, 44/*GG* i. pp. vii, 14–15; *GG* II. 147–148)—a continuation of the doubt expressed in (B) (*FA* 117). He may have continued to worry about the very unKantian notion of particular objects whose existence corresponds to and is justified in terms of no intuition. I believe, however, that intuitive, mathematical worries were more basic. Granting the notion of the extension of a concept, one can find certain instances of the principle to be obvious—instances that involve only quantification over individuals, or that use concepts that apply only to individuals. But once one considers quantification over (or concepts that apply to) higher-order objects—extensions of concepts—the principle is less obvious. Of course, this remark benefits from hindsight. It may thus seem anachronistic as an interpretation of Frege's intellectual situation.

[18] If the principle had been a law, the two sides of the biconditional would appear to meet a complex criterion for sameness of thought (sense) that Frege proposes in 1906 in a letter to Husserl (*PMC* 70/*WB* 105–106, Frege to Husserl, 11/10/1906). This criterion is, however, defective in several respects.

I do not think that it is. At the time he proposed Basic Law (V), Frege did lack acquaintance with Russell's paradox; so the specific danger for (V) of allowing indiscriminate quantification over higher-order objects could not have been apparent to him. On the other hand, his notion of the extension of a concept was intuitively unclear. It had previously been given little or no mathematical work. Frege explained the notion in terms of courses of values in 1891 (G & B 25, 30–32/*KS* 129, 133–135, "Funktion und Begriff", O 8, 15–19). And courses of values, especially those explained as objects derived from functions onto truth-values, were completely new to the mathematical scene (*BL* 6–7/*GG* i. 9–10). Uncertainty about Basic Law (V) must have derived partly from a felt unclarity about the notions of the extension of a concept and the course of values of a function. Frege was explicating the familiar notion of number, which he regarded as insufficiently determinate in its conventional mathematical employment, in terms that were substantially less familiar.[19]

[19] How *should* we understand the notions of the extension of a concept and the course of values of a function, as Frege uses them? I think that there is no clear answer. An illuminating one would require new mathematical work as well as philosophical explication. Frege provides an intuitive means of representing courses of values in terms of geometrical graphs, where the argument of the function is the numerical value of the abscissa and the value is the numerical value of the ordinate. (G & B 25, *KS* 129, "Funktion und Begriff", O 8; *GG* i. 129). No doubt this gives us some hold on the notion. We are then quite ready to explicate this notion of a graph in set-theoretic terms—as a set of ordered couples, where the members of the couples are the arguments and values of the function. This explication is, however, seriously at odds with Frege's own standpoint. The problem is not merely that Frege's notion of the course of values of a predicate lacks any limitations on the objects that can form members of the ordered couples. No consistent set theory is free from some such limitations; and one can hardly require a rationalized explication of Frege's notion to match it inconsistency for inconsistency. Rather, the point is that the particular limitations imposed by the dominant, iterative conception of sets are very unFregean. Frege conceived courses of values as being projections from predication, not constructions from elements (cf. sec. III). The idea of explicating a course of values, much less a *function*, in terms of ordered couples would have appalled him. From his point of view, a course of values' falling in its own extension is no more peculiar than a predicate's applying to itself. By contrast, a set's belonging to itself is ruled out from the outset by the iterative conception. Another way to put the point is that Frege's extensions of concepts violate the axiom of foundation. In fact, if the extensions of concepts are conceived as sets of ordered couples, not a single extension of a concept is well-founded in either the argument or the value position of the ordered couples. From the perspective of the iterative conception of set, this is tantamount, I think, to saying that courses of values are not sets at all.

There are, to be sure, set theories, often called "deviant," that violate the axiom of foundation. But I find it doubtful whether any of these provides an intuitively satisfying reconstruction of Frege's notion of courses of values (although making this point stick would require more argument than I am prepared to give here). The lambda calculus provides some explication of a Frege-like conception of function and concept—and even exhibits some ambivalence over the concept–object distinction. But current model-theoretic interpretations of the lambda calculus do not seem to me to be illuminatingly interpreted as reconstructing Frege's notion of the extension of a concept. Cf. D. S. Scott, "Lambda Calculus: Some Models, Some Philosophy", in J. Barwise, H. J. Keisler, and K. Kunen (eds.), *The Kleene Symposium*, (Amsterdam: North-Holland, 1980); and Albert R. Meyer, "What is a Model of the Lambda Calculus?" (xerox, Laboratory for Computer Science, MIT, Cambridge, Mass.). Perhaps this is just as well; or perhaps the situation will change. But as of now, it seems fair to say that Frege's notion has not found an intuitively attractive, mathematically fruitful counterpart.

For whatever reason, Frege himself did not find his notion and the associated axiom intuitively compelling. He seems to have struggled to clarify them between 1884 and 1893 with incomplete success. From Frege's own point of view, of course, the concept *extension of a concept* (or the relevant sense) could not itself be unclear (*FA* p. vii; *KS* 122, "Über das Trägheitsgesetz", O 158). In later work, Frege speaks of the expression as being defective and as lacking a *Bedeutung*. But how might he have rationalized the difficulty he was in before 1903?

The difficulty was that Frege found axiom (V) subject to doubt, but maintained that it expressed a logical law. On Frege's own view, logical laws were self-evident and undeniable without plunging thought into "complete confusion" (*BL* 127/*GG* ii. 251; *FA*, sec. 14). In 1891, and probably since 1884, the difficulty was even sharper: the two sides of the axiom were held to express the same sense—the same thought. Yet Frege harbored sufficiently strong doubts about the axiom to have considered radically different ways of founding arithmetic.

One may interpret Frege as lacking a coherent rationalization of the difficulty—as simply torn between the blandishments of the axiom and the warnings of his intuitive conscience. I think, however, that the rationalist elements in his epistemology that we noted in section I provided him with an explication of his situation. In writings early and late, Frege emphasizes that one may not clearly apprehend the concepts (senses) expressed by expressions one uses. Fully grasping the content or sense of an expression—and bringing its conventional significance up to the level of the sense it expresses—depended upon logical analysis. Such analysis involved refining the old language, or introducing new language, and embedding that language in a rigorous logical *theory* in such a way as to elicit thorough understanding of the truth conditions of, and logical relations among, senses expressed by sentences containing the refined language. Thus logical analysis was not separable from the acquisition of logico-mathematical knowledge. Frege thought that one attained insight into the relevant concepts or senses only through developing a theory and seeing it work. This rather pragmatic emphasis on the interdependence of theory and understanding is an integral part of Frege's rationalist conception. Frege retains the traditional rationalist insistence on the close relation between understanding or insight and *a priori* propositional knowledge. But he reverses the traditional order of priority. For him, full understanding depends on, or at most is co-equal with, knowledge, which derives from logico-mathematical theory. Frege's pragmatic emphasis occurs infrequently but repeatedly throughout his writings (cf. *PW* 33/*NS* 37, "Booles rechnende Logik und die Begriffsschrift" (1880–1881); *KS* 122–124, "Über das Trägheitsgesetz", O 157–161; *KS* 369, "Die Verneinung", O 150; *BL* 7, 25/*GG* i. pp. x, xxvi). Until a theory is developed and logical analysis fully carried out, a full grasp of the refined language may not be achieved.

The expression 'extension of a concept' was just such new or refined language. The attempt to interpret axiom (V) as a biconditional whose two sides express the same sense or thought, and indeed the mere postulation of the axiom as expressing a self-evident law of logic, would have been patently untenable *only* on the assumption that the sense of the refined language is perfectly grasped. It could not have surprised Frege that perfect mastery of the notion was difficult to attain (*FA* p. vii; *KS* 122, "Über das Trägheitsgesetz", O 158). He must have hoped that the law's self-evidence would emerge when the senses of the axiom and the expression 'extension of a concept' were fully grasped. (Further evidence for this speculation occurs in the method of *BL*, sec. 10.)

Frege's rationalist point of view is unquestionably strange from a contemporary perspective. The taproot of the strangeness is, I think, the conception of fully determinate senses or thoughts that are completely independent for their representational properties of human practices or understanding. This conception has as a corollary the conception of self-evident thoughts that no one understands. Few have found these conceptions palatable. I believe that the point of view has some strengths that survive this unpalatability. But here is not the place to explore them.

A common opinion regarding the turn-of-the-century paradoxes is that Law (V), and the closely associated naive comprehension axiom, were originally regarded as obvious—that their defeat at the hands of paradox came as a complete surprise. Surprise there was. But this interpretation underplays the degree of ferment and unclarity over foundational notions in the late nineteenth century. Analogs of the common opinion as applied to other mathematicians of the period need drastic revision. But the view fits even Frege poorly. The evidence indicates that he spent several years trying to dispense with his axiom and, alternatively, to provide it with an attractive exposition and intuitive justification. It is hard not to admire the persistence of his self-questioning.

8 *Frege on Knowing the Third Realm (1992)*

Anyone who reads Frege with moderate care is struck by a puzzle about the central objective of his work. His main project is to explain the foundations of arithmetic in such a way as to enable us to understand the nature of our knowledge of arithmetic. But he says very little about our knowledge of the foundations. A full treatment of this and associated puzzles would require more room than I have here.[1] I want to give a short solution to the puzzle, and then discuss one aspect of it that I find interesting.

The short solution is that Frege accepted the traditional rationalist account of knowledge of the relevant primitive truths, truths of logic. This account, which he associated with the Euclidean tradition, maintained that basic truths of geometry and logic are self-evident. Frege says on several occasions that such primitive truths—as well as basic rules of inference and certain relevant definitions—are self-evident. He did not develop these remarks because he thought they admitted little development. The interesting problems for him were finding and understanding the primitive truths, and showing how they,

I am indebted to Tom Ricketts for clarifying his views, discussed in note 15. I have also benefited from remarks from various participants at a conference on early analytical philosophy held at the University of Chicago in honour of Leonard Linsky.

The following works by Frege are cited in the text by the abbreviations that follow their titles: *The Basic Laws of Arithmetic* (*BL*), trans. and ed. M. Furth (Berkeley: University of California Press, 1967); *Begriffsschrift und andere Aufsätze* (*B*), ed. I. Angelelli (Hildesheim: Georg Olms, 1964; 2nd edn. 1977); *Foundations of Arithmetic* (*FA*), trans. J. L. Austin (Evanston, Ill.: Northwestern University Press, 1968; Oxford: Basil Blackwell, 1980); *Die Grundgesetze der Arithmetik* (*GG*) (Hildesheim: Georg Olms, 1962); *Kleine Schriften* (*KS*), ed. I. Angelelli (Hildesheim: Georg Olms, 1967); *Nachgelassene Schriften* (*NS*), ed. H. Hermes, F. Kambartel, and F. Kaulbach (Hamburg: Felix Meiner, 1968; 2nd edn. 1983); *Philosophical and Mathematical Correspondence* (*PMC*), trans. B. McGuinness and H. Kaal (Chicago: University of Chicago Press, 1980); *Posthumous Writings* (*PW*), ed. H. Hermes, F. Kambartel, and F. Kaulbach, trans. P. Long and R. White (Chicago: University of Chicago Press, 1979); *Wissenschaftlicher Briefwechsel* (WB), ed. G. Gabriel, H. Hermes, F. Kambartel, C. Thiel, and A. Veraart (Hamburg: Felix Meiner, 1976); *Collected Papers on Mathematics, Logic, and Philosophy* (*CP*), ed. B. McGuinness (Oxford: Basil Blackwell, 1984). The sign "O" marks the pagination of the original publications of Frege's articles.

[1] An auxiliary puzzle attends this primary one. Most of Frege's philosophical work is directed at correcting what he regards as the misunderstandings embedded in normal practice and language—misunderstandings that he thought had prevented a correct understanding of the fundamental notions present in his account of the foundations. But he has even less to say about the epistemology of his analysis and elucidation of the notions that interested him than he does about knowledge of the foundations.

together with inference rules and definitions, could be used to derive the truths of arithmetic.

This short solution seems to me correct—as far as it goes. It does, however, leave out a lot. Frege thought that knowledge of the axioms of geometry require intuition—an imaginative or broadly perceptual capacity (*FA* pp. 19–21). Knowledge of the basic truths of logic simply required reason. He regarded both types of basic truths as self-evident, but the differences between the two types of knowledge are significant. That is one complication. Another is that Frege uses a variety of terms that are translated "self-evident". His sophisticated understanding of the notion is neither psychologistic nor purely proof-theoretic. He does not mean by it what most contemporary philosophers would mean by it. His uses of it relate in interesting ways to his basic philosophical views. A third complication is that there are complex relations between Frege's appeals to self-evidence and an appeal he makes to pragmatic epistemological considerations. This appeal makes his rationalism original and gives it, I think, special relevance to modern problems. Although these points are worth developing, I will not discuss them here. Instead, I shall discuss an intensification of the puzzle in the light of the short solution that I have just given.

Frege assumes that only truths are self-evident. He also assumes that it is rational to believe what is self-evident, given that it is well understood. Frege believes in other types of purely mathematical justification for arithmetical judgments besides self-evidence and derivation from self-evident truths.[2] But these other types also involve only reason. The key idea in what follows is that Frege assumes that we can know arithmetic and its foundations purely through reason, and that individuals are reasonable and justified in believing basic foundational truths (e.g. *PW* 175; *NS* 190, "17 Kernsätze zur Logik" (*c*. 1906)).

Frege held that both the thought contents that constitute the proof-structure of mathematics and the subject matter of these thought contents (extensions, functions) exist. He also thought that these entities are non-spatial, non-temporal, causally inert, and independent for their existence and natures from any person's thinking them or thinking about them. Frege proposed a picturesque metaphor of thought contents as existing in a "third realm". This "realm" counted as "third" because it was comparable to but different from the realm of physical objects and the realm of mental entities. I think that Frege held, in the main body of his career, that not only thought contents, but numbers and functions are members of this third realm.[3] (Cf. *FA* p. viii; *BL*

[2] I distinguish purely mathematical justifications from justifications of mathematics that derive from applications to the empirical world—which he also seems to have believed in, but which I lay aside.

[3] Frege's logic is not committed to thought contents, only to extensions and functions. But this is an artifact not of his views about logic, but of his interests in deriving arithmetic from logic. For that, he did not need to refer to thought contents (*Gedanken*). But he clearly envisioned a logic committed to

15–16; *GG* p. xvii). Entities in the other realms depended for determinate identities on functions (concepts) in the third realm. Since logic was committed to this realm, and since all sciences contained logic, all sciences were committed to and were partly about elements of this realm. Broadly speaking, Frege was a Platonist about logical objects (like numbers and truth-values), functions, and thought contents. I shall say more about Frege's Platonism later, but I think that I have said enough to enable me to introduce the problem that I want to discuss.

The problem is that of understanding how reason alone could justify one in believing that a thought is true, when the thought has a subject matter that is as independent of anyone's thinking as Frege indicates it is. How could mere reasoning give one any ground for believing that a realm of entities is one way rather than another, when that realm is so independent of that reasoning? How could reasoning and understanding have any tendency to tell one how things in such a realm really are?

This problem is clearly kin to a problem about the relation between the knowledge and truth of mathematics that is commonly discussed today.[4] The contemporary problem is that of understanding how our beliefs about mathematics could have any tendency to be true, given that we do not appear to bear causal-perceptual relations to the subject matter of mathematics. This may be seen as a problem for Frege. But it is not one that he would have naturally formulated for himself. His attitude toward the point that numbers and thought contents are not causally effective ("wirklich") seems to have been "so what?".[5] He showed no special interest in the causal theory of knowledge, or in cashing out his occasional physical-contact metaphors of "grasping" thoughts. The idea that mathematical or logical knowledge should be judged by reference to the standard of empirical knowledge would have seemed foreign to him.

Like Frege I see no reason to think that mathematical or logical knowledge is questionable because it apparently lacks causal-perceptual relations to its subject matter. But I formulated a problem that makes no reference to causal-perceptual relations. This formulation seems not to import assumptions foreign to Frege. A theory of knowledge should not make it puzzling how being reasonable could be conducive to having true beliefs. Frege's rationalist

thought contents. In the correspondence with Russell, for example, he indicates the need for special names of senses to avoid the "ambiguity" of indirect discourse or propositional attitude attributions (cf. *PMC* 153; *WB* 236, Frege to Russell, 12/28/1902).

[4] Paul Benacerraf, "Mathematical Truth", in P. Benacerraf and H. Putnam (eds.), *Philosophy of Mathematics* (Cambridge: Cambridge University Press, 1983).

[5] Actually, he does provide an argument: objective sense perception requires perceptual belief; but perceptual belief requires grasp of thoughts in the third realm—a non-causal relation; so one cannot cite the element of causal interaction in sense perception as providing grounds for thinking that knowledge cannot involve non-causal cognitive relations to abstract entities (*CP* 369–70; *KS* 360, "Der Gedanke", O 75).

theory of knowledge combines with his Platonism to raise a question at just this point. Why did he not discuss the question?

Some recent interpretations of Frege suggest that it is a question that is somehow precluded by his philosophy, or that it rests on fundamental misreadings of his views. One might question the notion of "subject matter" that the formulation of the problem uses. Or one might claim that Frege's notion of truth or of logic blocks a "meta-standpoint" from which one could raise the question. Or one could doubt whether Frege's Platonism should be understood in the way that the "third realm" metaphor suggests, and maintain that in talking about numbers or thought contents, Frege was really talking about our language or our cognitive practices in such a way that no gap between our beliefs and the numbers was even formulable. I will not criticize in detail all such lines for short-circuiting our question for Frege, though I will remark on some of them in a general way. I think none provides good grounds for ignoring the question. In fact, Frege himself gives an answer to it. The reason why he did not discuss it in detail is similar to the reason why he did not discuss knowledge of the foundations in detail. He believed that he had little to add to a traditional answer. I think that his answer is worth understanding.

Let us back up a bit. I want to explain in more detail what I mean by saying that Frege was a Platonist about logical objects, functions, and thought contents.

First, some preliminary disclaimers. Although I think that Frege maintained a metaphysical view about numbers and other such entities, I do not believe that this view dominated his thinking. His is, for the most part, the relaxed Platonism of a mathematician who simply assumes that there are numbers, functions, and so on, and who regards these as an abstract subject matter which can be accepted without special philosophical explanation, which is clearly different from mental or physical subject matters, and which mathematics seeks to characterize correctly. One can see this attitude toward functions very prominently in "On Function and Concept". Frege highlighted the intersubjective objectivity of scientific theorizing. He believed that standard mathematical practice tells one most of what is true about mathematical entities, and he thought that one could know mathematical truths independently of any philosophy. Indeed, he assumes that ordinary mathematical practice yields *certain* knowledge even prior to the execution of his foundationalist program (*B* §13; *FA* §2).

Most of Frege's uses of his metaphysical view are defensive. His metaphysical remarks ward off idealist, physicalistic, psychologistic, reductive, or deflationary positions because he thinks that they prevent clear understanding of the fundamental notions of logic and arithmetic. As I shall later show, he does give his Platonism extra-mathematical work. But he does not think out this side of his philosophy as someone would who was concerned about certainty or who believed that logic and mathematics has no other cognitive underpinning than that provided by philosophy.

Another preliminary point about Frege's Platonism is that although he uses the Platonic metaphor of vision on occasion, he shows no interest in developing the metaphor when characterizing our knowledge. He appeals to no faculty other than reason in his account of our mathematical knowledge. Moreover, as I have intimated earlier, his epistemological views are complex, and involve not only Platonic elements, but elements not at all associated with traditional Platonism. The discussion in what immediately follows will be concerned with the Platonic character of Frege's *ontology*. For now, I lay epistemology aside.

As is well-known, Frege thought that extensions—including numbers—functions—including concepts—and thought contents are imperceptible, non-spatial, atemporal, and causally inert.[6] He emphasizes that numbers (*FA* p. 108), concepts (*FA* p. vii), and thought contents (*BL* 23; *GG* p. xxiv) are discovered—not created. He sharply distinguishes the act of thinking, which does occur in time, from the thought contents that we "grasp" or think, which are timeless. So in coming to know thought contents that denote numbers, concepts, and the like, one discovers objects, concepts, and relations that are what they are timelessly, independently of any causal influence. One comes to "stand in relation", as Frege says, with non-spatial, atemporal entities (*CP* 363, 369; *KS*, 353–4, 360, "Der Gedanke", O 60–61, 66; *BL* 23; *GG* p. xxiv.)

Frege calls numbers, concepts, and thought contents "objective". By this he means, partly, that they are not intrinsically borne by a mind, as a pain or an afterimage is. He says that they are subject to laws. They are common property to different rational beings (*FA* §26; *CP* 363 ff.; *KS* 355 ff., "Der Gedanke", O 69 ff.). Much of Frege's discussion of atemporal entities centers on their objectivity. For many of his purposes, the intersubjectivity and lawfulness of logic are its key properties.

Many of these things might be maintained by someone who is not a Platonist. One might make the remarks about imperceptibility, non-spatiality, atemporality, and causal inertness, if one glossed them as part of a practical recommendation or stipulation for a theoretical framework, having no cognitive import—or as otherwise not being theoretical claims or claims of reason. Carnap might have said at least some of those things, though only

[6] Numbers are counted imperceptible (*FA* §85; *PW* 265; *NS* 284, "Zahl" (1924)). Thoughts are termed imperceptible (*CP* 369; *KS* 360, "Der Gedanke", O 75). Numbers are counted non-spatial (FA §§58, 61, 85, 93). Thoughts are counted non-spatial (*CP* 369–370; *KS* 360, "Der Gedanke", O 75). Concepts or other functions are counted atemporal and by implication imperceptible, non-spatial, and causally inert (*FA* pp. vii, 37; *CP* 133; *KS* 122, "Über das Trägheitsgesetz", O 158). He also suggests these points about concepts indirectly (*BL* 23; *GG* p. xxiv; *CP* 198; *KS* 181–182, "Rezension von: E. G. Husserl, *Philosophie der Arithmetik* i.", O 317–318). Numbers are counted atemporal (*CP* 230; *KS* 212, "Le Nombre entier", O 74). Thoughts are counted atemporal (*CP* 379–370; *KS* 360, "Der Gedanke", O 75). Numbers are counted causally inert (*FA* §85; *BL* 15–16; *GG* p. xviii). Thoughts are said to be causally inert (*BL* 23; *GG* p. xxiv; *PW* 137–138; *NS* 149–150, "Logik" (1897); *CP* 230, 371; *KS* 212, "Le Nombre entier", O 74; *KS* 361–362, "Der Gedanke", O 76–77).

given certain background qualifications. Or one might have some other basis for qualifying these remarks, reading them as "non-metaphysical" or as lacking their apparent ontological import. Moreover, certain idealists might say these things. Kant might have said them, given certain background qualifications. He could have seen numbers as just as genuine and discoverable as physical objects are. And he could see their objective status in terms of the possibility of intersubjective agreement on laws governing them. Platonism has no monopoly on claims to lawlike or intersubjective objectivity about non-spatial, atemporal entities. So we need to say more in order to distinguish Frege's view from alternatives.

I would not take very seriously a reading of Frege as a Carnapian. Discussing my attitude would require going more into his methodology and epistemology than I plan to. I think it clear, however, that Frege was trying to provide a rational foundation for mathematics—in a way that Carnap would have regarded as misguided. Frege saw reason, not practical recommendation, as giving logical objects to us (e.g. *FA* §105). There is nothing remotely akin to Carnap's Principle of Tolerance either in Frege's philosophical pronouncements, or even more emphatically, in his temperament.

What interests me more is the distinction between Frege's Platonism, on the one hand, and certain idealisms or certain vaguer "practice" oriented anti-Platonisms, on the other. Platonism, as I understand the doctrine, regards some entities (for Frege, some objects and all functions) as existing, or being otherwise "real", non-spatially and atemporally. Further, it avoids commenting on them as having special status, including being dependent for their existence or nature (as opposed to their discovery) on practice or mental activity. They are in *no way* derivative, instrumental, fictional, or otherwise second-class. The relevant entities are fundamental. It would be incompatible with Platonism to regard them as essentially part of an appearance or perspective for a thinker—as Kant would have—though they may impose constitutive conditions on such appearances or perspectives. Platonism rejects any deeper philosophical commentary that would indicate that the nature or existence of these atemporal entities is to be regarded as in any way dependent on something mental, linguistic, communal, or on anything like a practice or activity that occurs in time. In Kant, we find a non-Platonic explanation of mathematical structures in terms of a mental activity, "synthesis", that underlies the categories and the forms of spatial and temporal intuition. And in Hegel abstract structures are held to be abstractions from spirit in history. Recently, some philosophers have sought to avoid being "metaphysical", contenting themselves with generalized remarks that mathematical objects are grounded in some unspecified way in linguistic or mathematical practice. Such views can admit non-spatio-temporal entities and can grant them objective status. But they are not Platonic in my sense. They regard atemporal entities as derivative from human practices—such as linguistic activity. I see such views as covertly idealist. Idealism regards actual activity or practice as implicated in the nature and existence of non-spatio-temporal structures. Platonism holds that structure is more fundamental than actual activity.

Frege's Platonism shows itself in two ways. One is that he never enters the commentary that an idealist (or a deflationist) would enter on his claims about non-spatio-temporal entities, or about their objectivity or their discoverability. He takes them to be fundamental. The other is that he claims, more than once, that the assumption of the relevant entities explains the intersubjective objectivity of science and communication. I will discuss these points briefly, in turn.

There is, as far as I can see, no evidence that Frege thought that the existence or nature of these non-spatio-temporal entities was to be explained in terms of human language, human inference, human practices (including the *activity* of judgment), or other patterns of human activity in time. Frege thought of extensions, functions, and thought contents as genuinely existing entities.[7] He opposed thinking of such entities as having some derivative status. He inveighs against any suggestion that they are products of the mind, mere symbols, or otherwise dependent on events in time.[8] Had he maintained that extensions, functions, or thought contents were dependent on human conceptualization or human language, judgment, or inference (actual or possible), he would have said so, and thereby qualified the numerous remarks that have traditionally invited the Platonic interpretation of his work. He never does say so. His claims that atemporal entities are independent of us are unqualified.

On several occasions, Frege compares the objectivity and existence of numbers, concepts, or thought contents with the existence and objectivity of physical objects. He compares numbers to the North Sea as regards objectivity (*FA* p. 34). In doing so, he very explicitly indicates that the entity that we call "the North Sea" is what it is completely independently of our imposing the boundaries or making a map that we use to associate that entity with the name "the North Sea". He elaborates this comparison elsewhere:

> Just as the geographer does not create a sea when he draws boundary lines and says: the part of the ocean's surface bounded by these lines I am going to call the Yellow Sea, so too the mathematician cannot really create anything by his defining. Nor can one by pure definition magically conjure into a thing a property that in fact it does not possess—save that of now being called by the name with which one has named it. (*BL* 11; *GG* p. xiii)

[7] He quantified over them with quantifiers of different types. He used first-order quantifiers for the objects, second-order quantifiers for the functions. The quantifiers are appropriately read as involving existential commitments.

[8] For extensions and numbers, cf. *FA*, *passim*; *BL* 10, 12; *GG*, pp. xiii, xiv; *BL* 15–16; *GG*, p. xviii; *CP*, 230; *KS* 212, "Le Nombre Entier", O 74. For concepts or functions, cf. *FA*, p. viii; *CP*, 133; *KS*, 122, "Über das Trägheitsgesetz" O 158. For thoughts, cf. *CP*, 363, 370; *KS* 353–354, 360–361, "Der Gedanke", O 60–61, 66–67.

He compares a mathematician's relation to numbers with the astronomer's relation to the sun (*PW* 7; *NS* 7, "Logik" (zwischen 1879 und 1891)) and to the planets (*FA* p. 37). He says that like geographers, mathematicians cannot create, but can only discover "what is there and give it a name" (*FA* p. 108; cf. also *BL* 23–24; *GG*, p. xxiv; *PW* 137; *NS* 149, "Logik" (1897)). He compares our epistemic relation to numbers and concepts (and probably thought contents) to our grasping a pencil:

> The picture of grasping is very well suited to elucidate the matter. If I grasp a pencil, many different events take place in my body . . . but the pencil exists independently of them. And it is essential for grasping that something be there which is grasped . . . In the same way, that which we grasp with the mind also exists independently of this activity . . . and it is neither identical with the totality of these events nor created by it as a part of our own mental life. (*BL* 23–24; *GG*, p. xxiv; cf. *PW* 137; *NS* 149, "Logik" (1897))

Thought contents exist independently of thinking "*in the same way*", he says, that a pencil exists independently of grasping it. (The artifactual character of pencils plays no role in his understanding of the analogy, as other examples indicate.) He says that thought contents are true and bear their relations to one another (and presumably to what they are about) independently of anyone's thinking these thought contents—"just as a planet, even before anyone saw it, was in interaction with other planets" (*CP* 363; *KS* 354, "Der Gedanke," O 69). And he compares a thought's independence of our grasping it to the star Algol's independence of anyone's being aware of it (*CP* 369; *KS* 359, "Der Gedanke", O 75).

All these comparisons suggest (and those of *BL* 23–24; *GG*, p. xxiv; *CP* 363, 369; *KS* 354, 359, "Der Gedanke," O 69, 75, explicitly state) that numbers, functions, and thought contents are independent of thinkers "in the same way" that physical objects are. Frege nowhere asserts or clearly implies that he maintains any sort of idealism—Kantian or otherwise—about the physical objects studied by the physical sciences. He nowhere qualifies the ontological status of physical objects. It is dubious historical methodology to attribute to a philosopher with writings that stretch over decades, a large, controversial doctrine, if he nowhere clearly states it in his writings. If Frege had believed in any such idealism about physical objects (or any doctrine qualifying their ontological status), he would have surely said he did.[9] Doing

[9] Such passages as *PW* 137; *NS* 149, "Logik" (1897), or any of the various passages about independence of mind that I discuss below, would require strong qualification, which Frege nowhere makes, to be compatible with any sort of idealism or deflationary reading. For an interpretation of Frege as a Kantian idealist, see Hans Sluga, *Gottlob Frege* (London: Routledge and Kegan Paul, 1980), e.g. 59–60, 115–116. Sluga cites mainly considerations that are external to Frege's texts. He also writes, "the central role of the Fregean belief in the primacy of judgments over concepts would seem to be explicable only in the context of a Kantian point of view". Sluga does not explain this remark. I think it misleading. Judgments and inferences are a source of discovery. But logical theory is about the forms of correct judgment and inference—not about judgments and inferences. Frege regards judgment as a form. (Cf. *CP* 383–385; *KS* 372–374, "Die Verneinung," O 152–154.) I know

so would have been necessary for a philosopher to balance the flat-out statements about mind-independence that Frege makes.[10]

Frege thought that to know the physical world, one has to grasp thoughts (which bore for him eternal denotational relations to concepts and extensions) that are eternal and eternally true. Logic is embedded in the content of any knowledge. Since logic is about (denotes) concepts and other functions, relations, and logical objects, all knowledge is at least partly about non-spatio-temporal entities. Moreover, logic concerns the forms of correct judgment and inference; and logical structure is discovered by reflecting on patterns of correct judgment and inference. But Frege does not give the slightest indication that he thought that either the physical world or the non-spatio-temporal entities inevitably appealed to in knowing it depend in any way on any activities of judgment, inference, or linguistic practice.[11]

Frege not only compares non-spatio-temporal entities to physical objects in their independence of us; he makes unqualified statements about the

of no evidence that he saw this form as dependent for its nature or existence on actual activities of judgment, or on anything like Kantian synthesis; there is substantial evidence that he did not.

[10] Some philosophers have suggested that Frege's use of the context principle somehow suggests a qualification on his Platonism. Issues surrounding Frege's context principle(s) are, of course, extremely subtle and complex. But it seems to me that the suggestion must involve some confusion. The context principles govern relations between linguistic expressions and their senses or referents. They do not bear directly on the nature of the senses or referents themselves at all. At most one of the principles might be coherently thought to rule out certain naive forms of epistemological Platonism (those that require that we have perception-like intuition of mathematical objects). There are many complex issues here, and some of them are not completely independent of ontology. But I think that any simple appeal to the context principles to motivate opposition to my interpretation will confuse language and epistemology with ontology.

[11] An interpretation of Frege similar to Sluga's is proposed in Joan Weiner, *Frege in Perspective* (Ithaca, NY: Cornell University Press, 1990). In interpreting the North Sea comparison (*FA* 34), Weiner notes Frege's remark that if we should happen to draw the boundaries of what we call "the North Sea" differently, what we now call "the North Sea" would still exist, though perhaps unrecognized. But she continues, "It is important to realize, however, that the claim that such unrecognized objects exist need not be a substantive metaphysical claim. For . . . to claim that unrecognized objects exist is simply to claim that it is possible to formulate (heretofore unformulated) concepts under which exactly one object falls" (p. 171). Weiner cites no texts to support this reading. I see no reason to think that existence claims for Frege are "simply" claims about possibility or about formulations; he gives every indication that they are not about possibility, language, or activity at all. Later she correctly claims that Frege believed our knowledge requires language or drawing boundaries. But she moves without argument from this remark about knowledge to one about the world: "Frege's view is that the physical world is not articulated—that we impose structure on it" (p. 184). The language of imposition is not present or implied in Frege. That concepts mark boundaries of the ocean is nowhere said to depend in any way on anything about language or people. (Similarly, with concepts demarcating possible numerations in such cases as packs of cards.) Frege writes: "To bring an object under a concept is merely to recognize a relation that already existed beforehand" (*CP* 198; *KS* 181–182, "Rezension von: E. G. Husserl, *Philosophie der Arithmetik* i.", O 317; cf. *PW* 137; *NS* 149, "Logik" (1897)). Weiner glosses Frege's claims that mathematical truths are independent of us by *excepting* an alleged presupposed need to impose structure and formulate boundaries linguistically (pp. 201 ff.). She further writes, "discovering what is 'there' in the 'realm of the abstract' amounts to discovering what meets the descriptions that interest us" (p. 203). Weiner cites no texts to support either of these claims. Frege makes no exceptions or qualifications on his claims of independence; he notes no such presupposition. And it is at best deeply misleading to say that for Frege discovering mathematical structures "amounts to" discovering something associated with words, our interests, or ourselves.

independence of such entities from anything about us. He repeatedly claims that both the truth of thought contents and thought contents themselves are independent of individuals' and groups' thinking the thoughts or recognizing them to be true (*BL* 15, 23; *GG*, pp. xvii, xxvi; *FA* p. 60; *CP* 363; *KS* 354, "Der Gedanke", O 69). He writes: "What we want to assert in using that proposition [that the number three is prime] is something that always was and always will be objectively true, quite independently of our waking or sleeping, life or death, and irrespective of whether there were or will be other beings who recognize or fail to recognize this truth" (*CP* 134; *KS* 123, "Über das Trägheitsgesetz", O 159).

The lack of qualification in his claims of independence is especially striking in two passages: one where he writes that someone's thinking a thought has "nothing to do" either with its truth or with the thought content itself (*CP* 368; *KS* 359, "Der Gedanke", O 74); and another where he writes that thought contents are not only true independently of our recognizing them to be true, but they, the thought contents themselves, are "absolutely independent of our thinking" (*PW* 133; *NS* 145, "Logik" (1897)). Independence is independence. Frege's repeated remarks about mind-independence of non-spatio-temporal entities would not have been literally true, if they had been backed by a set of unstated qualifications of the sort that an idealist (or deflationary) interpretation of them would require. Ultimately the idealist asserts dependence of the thought contents and timeless objects on some underlying practice or activity that makes possible the framework in which attributions of objectivity are made. No idealist—and no deflationist who thinks that non-spatio-temporal entities are dependent on our language, practices, or judgments, or who thinks that general philosophical assertions about them are "non-factual"—would have made such statements without careful, explicit qualifications. Frege enters no such qualifications.

Frege repeatedly inveighs against seeing logic (or mathematics) as embedded in language in the way that grammar is.[12] He thought that thought contents, logical objects, and logical functions bear no such essential dependence relation to the actual practice of thinking or language use. For Frege, the

When our interests and descriptions happen to accord with mathematical truth, we do, of course, discover things that "meet" those interests and descriptions. But Frege explicitly says that our relation to logical truths and mathematical structures is "inessential" to their nature and existence (*CP* 371; *KS* 361, "Der Gedanke", O 76).

[12] Some interpreters of Frege have taken his views to be redescriptions of features of our practices of judgment or of linguistic use. Although Frege does describe logical structures that inform linguistic and cognitive practice, and does think that by reflecting on and reforming such practice we can discover these structures, I know of no evidence that Frege thought that the theory of judgment is really fundamentally about the activity or practice of judgment, much less linguistic practice. It is important to distinguish Frege's method of discovery (which does focus on language and activities of judgment) from Frege's views about the nature of thought contents and of judgment. It is also crucially important to realize that Frege was interested in judgment as a norm-yielding form, not in judgment as a human activity. Frege thought that thought contents and the form of judgment bear no essential relation to either language or activities (practices) of judgment, potential or actual, of human beings.

subject matter of logic is not the nature of human thinking or practice (*BL* 13; *GG* p. xvi), *even when that practice accords with the laws of truth*: "But above all we should be wary of the view that it is the task of logic to investigate actual thinking and judging, insofar as it is in agreement with the laws of truth" (*PW* 146; *NS* 158, "Logik" (1897)—the published translation is ambiguous in a way that does not match the German). This independence insures, for Frege, no scope for variation in the laws of logic between one group of thinkers and another (*PW* 7, 146; *NS* 7, "Logik" (zwischen 1879 und 1891); *NS* 158, "Logik" (1897); *BL* 13; *GG* pp. xv–xvi).

Frege criticizes one Achelis who writes,

> ... the norms which hold in general for thinking and acting cannot be arrived at by the one-sided exercise of pure deductive abstraction alone; what is required is an empirico-critical determination of the objective principles of our psycho-physical organization which are valid at all times for the great consciousness of mankind.

Frege replies:

> [It appears that according to Achelis] the laws in accordance with which judgments are made are set up as a norm for how judgments are to be made. But why do we need to do this? ... Now what is our justification for isolating a part of the entire corpus of laws and setting it up as a norm? ... Are [the laws of logic] like the grammar of a language which may, of course, change with the passage of time? This is a possibility we really have to face up to if we hold that the laws of logic derive their authority from a source similar to that of the laws of grammar ... if it is normal to judge in accordance with our laws of logic as it is normal to walk upright. (*PW* 147; *NS* 159, "Logik" (1897))

Frege thought one can discover logic by reflecting on linguistic and mathematical practice. But he makes it very clear that his logical theory is not about practices, and does not take its authority from such practices. They are not what ground the normative structures that logic articulates.[13]

A second way Frege's Platonism shows itself lies in his attempts to explicate the success of science, the fact of intersubjectively objective cognitive practices, and indeed the authority of logic, in terms of the timelessness of the truths and structures of logic. In *The Basic Laws of Arithmetic* he states that the laws of logic (which he also calls the laws of truth) are authoritative *because* of their timelessness: "[the laws of truth] are boundary stones set in an eternal foundation, which our thought can overflow, but never displace. It

[13] Frege writes that there is no contradiction between something's being true and everyone's taking it to be false (*BL* 13; *GG* pp. xv–xvi), making it clear that he does not believe in some general connection between thought contents (or intentional contents, or what are expressed in language) and actual judgments and practice, that would close any possible gap between mind and subject matter. There is more evidence for this fact in his discussion of scepticism in "The Thought". The example he gives in *BL* 13/*GG* pp. xv–xvi concerns an empirical truth. As I discuss below, he indicates that some truths—simple truths of arithmetic and basic logical truths—can be denied only through madness, and that any attempt to deny them in a thoroughgoing way will undermine judgment itself.

is because of this that they have authority for our thought if it would attain to truth" (*BL* 13; *GG* p. xvi).

Frege frequently claims that only because individuals are not the bearers of thought contents is scientific communication possible (*BL* 17; *GG* p. xix; *PW* 133, 137–138; *NS* 144, 149–150, "Logik" (1897); *CP* 362, 368; *KS* 353, 359, "Der Gedanke", O 69, 74). But he sometimes goes further. In the Logic manuscript of 1897 he indicates that it is the timelessness of the subject matter of logic—the laws' not containing conditions that might be satisfied at some times and not at others—that enables logic to provide completely universal laws of truth (*PW* 148; *NS* 160). The idea seems to be that all true thoughts are eternally true if they are true at all; but some have temporal subject matters. Some true laws even contain conditions that might be satisfied at certain times but not at others. But the laws of logic cannot be about temporal subject matters and cannot contain such conditions. For if the truth of some thought follows from the truth of others, then it must always follow. So to account for the universal aspect of entailment, one must assume that the subject matter of logic is eternal. (The conclusion of this argument, though not the argument itself is stated in *BL* 13; *GG* p. xvi.)

In "Thoughts" Frege gives two more arguments that scientific objectivity (of communication and of knowledge of the physical world, respectively) is explicable only on the view that thought contents belong to a "third realm" that is neither mental nor physical. In the first argument (*CP* 362; *KS* 353, "Der Gedanke", O 69) he holds that scientific communication cannot be understood on the assumption that thought contents are ideas in particular people's minds. He had previously maintained that thought contents are clearly not perceptible or knowable on the basis of perceptions (*CP* 354, 360; *KS* 345, 351, "Der Gedanke", O 61, 66). He concludes (*CP* 362; *KS* 353, "Der Gedanke", O 69) that in order to understand the objectivity of the communal scientific enterprise, one "must recognize" the third realm. The timelessness of the truths of this realm and the fact that their truth is independent of whether anyone takes them to be true are clearly seen as part of an account of how a "science common to many on which many could work" is possible.[14]

In the second argument (*CP* 368; *KS* 359, "Der Gedanke", O 74), Frege indicates that a "firm foundation of science" must be facts that are independent of men's varying states of consciousness. Facts are, he maintains, true thoughts. True thoughts have the requisite independence: not only are they not part of anyone's "inner" mental world; their truth "has nothing to do with" someone's thinking them. The work of science consists in the discovery of true thoughts (which provide a "firm foundation" for the science).

[14] Although this argument is not explicit in *FA*, the attitude behind it is not hard to discern in the introduction (*FA* pp. vii–viii). I think that the argument is the least interesting of the three arguments I am discussing.

Moreover, Frege argues that the applicability of mathematical truths to investigations at any time (he cites application of mathematics in an astronomical investigation into events in the distant past) is possible *because* a mathematical thought's truth, and the thought content itself, are timeless. So, he concludes, explicating the objectivity of science and the temporally neutral applicability of mathematics requires that both the thoughts and their truth be timeless.[15]

These arguments take for granted the existence of the objectivity manifested in intersubjectively accepted norms for communication and the checking and confirming of scientific results. Frege thinks that we need no reassurance about its solidity. He is not concerned with scepticism. He regards ordinary certainties as certain (*B* §13; *FA* p. 2). He does not seek foundations, nor does he appeal to his Platonism, to bolster confidence in an otherwise doubtful scientific enterprise. He does not view philosophy in the grand manner, as protecting science against otherwise dangerous philosophical worries. He articulates his Platonism because he finds that a refusal to qualify the timelessness of mathematical structures, or to explain them in

[15] Thomas G. Ricketts, "Objectivity and Objecthood: Frege's Metaphysics of Judgement", in *Frege Synthesized* (Dordrecht: Reidel, 1986) opposes reading Frege as "the archetypical metaphysical platonist" (p. 65), according to which "the mind-independent existence of things is for Frege a presupposition of the representational operation of language: it explains how our statements are determinately true or false apart from our ability to make or understand them". This description of Platonism does not fit the Platonism I attribute. Frege was clearly not trying to give a general explanation of linguistic representation or even of intentionality in judgment. But—in contrast to Ricketts—I think that Frege thought, as the previously cited passages indicate, that assuming the mind-independence of all thought contents, concepts, and logical objects, is necessary to understanding the objectivity of scientific practice and the universal applicability of logic and mathematics. I do not think that he thought that such objectivity would somehow be in jeopardy in the absence of such an explanation. Logic was for him epistemically prior to philosophy of logic. It is rather that such an explanation accounts for what is involved in judgment, logical inference, and logical truth. Ricketts elaborates: "The crucial feature of this [Platonist] line of interpretation is its taking ontological notions, especially that of an independently existing thing, as prior to and available apart from logical ones, from notions of judgment, assertion, inference, and truth" (p. 66). Ricketts also thinks that Frege's claims about the objectivity associated with judgment are not meant to be factual claims, and that there is therefore no possible explanation for Frege of objectivity. As I indicate in the text, I see no evidence for a relevantly applicable distinction in Frege between factual and non-factual claims. Moreover, Frege's Platonism does not involve any claim about the priority of ontological notions over logical notions. (I do not see even initial plausibility to attributing this assumption to Frege.) Logic and ontology are mutually entangled in Frege. Logic is about what is, as Frege says (*CP* 351; *KS* 342, "Der Gedanke", O 58); it has an ontology. But logic is the most general science. So no thought about being could be independent of its notions. Moreover, Frege's most fundamental ontological categories (function and object) are logical categories. Nor does Frege's appeal to ontology in his account of the objectivity of science and the universality of logic imply that he thought that ontological notions are prior to logical ones, much less available apart from them. The explanation is not a definition, derivation, or reduction. All the key ontological notions he uses both presuppose and include logical notions. Rather he thought that a full understanding of logic involves appeal to notions like logical object, function, thought content, mind-independence, timelessness, causal inertness, imperceptibility, non-spatiality. Frege thought of himself as describing the ontological features that logic must have. Logical and ontological notions are interrelated for Frege; and all the relevant logical objects and functions are timelessly related to the relevant notions (thought components). Frege sees the whole logical structure, not just objects, in a Platonic fashion.

terms of something more familiar and temporal—such as our minds or practices—provides the best understanding of scientific intersubjective objectivity. He thinks his view shows why practices that have been found to be firm are in fact firm (*FA* p. 2).

Let us turn to Frege's views about how we know this third realm of entities. As I indicated earlier, I am prescinding from complexities in Frege's epistemology. What is important for our purposes is that Frege thought that our knowledge of the primitive logical truths and inference rules depends on a logical faculty—reason (*FA* p. 21; *PMC* 37, 57; *WB* 63, Frege to Hilbert, 12/27/1899; *WB* 89, Frege to Huntington undated; *GG* ii. §74; *CP* 405; *KS* 393, "Gedankengefüge", O 50; *PW* 267–273; *NS* 286–292, "Erkenntnisquellen der Mathematik und der mathematischen Naturwissenschaften" (1924/1925)). The question is: how could Frege believe that reason alone could give one knowledge of an atemporal realm of entities that are completely independent for their existence, nature, and relations to one another, of anyone's reasoning?

Frege is aware that foundational questions about our knowledge of mathematical structures ultimately come down to questions about knowledge of the primitive truths and inference rules. He is admirably clear that logic does not answer these questions: "The question why and with what right we acknowledge a law of logic to be true, logic can answer only by reducing it to another law of logic. Where that is not possible, logic can give no answer" (*BL* 15; *GG* p. xvii); "We justify a judgment either by going back to truths that have been recognized already or without having recourse to other judgments. Only the first case, inference, is the concern of Logic" (*PW* 175; *NS* 190, "17 Kernsätze zur Logik" (*c*. 1906)). Frege lays fundamental epistemological questions aside in much of his work, especially in *Basic Laws of Arithmetic*. But it would be a serious misunderstanding to think that he thought that the questions were off limits.[16] For he expresses a consistent interest in them.

Of course, he thought that one could not and need not argue for the basic logical truths. But he did see them as a source for the justification of the belief in them by a person who understands them. He thought that they are self-evident. We justify our judgment of the basic truths, as he said, without having recourse to other judgments (*PW* 175; *NS* 190, "17 Kernsätze zur Logik" (*c*. 1906)).

One needs to bear in mind here a three-fold distinction that Frege often carries along in his writings (it is very explicit in "The Thought", *CP* 352; *KS*

[16] Cf. also *PW* 3, 175; *NS* 3, "Logik" (zwischen 1879 und 1891); 190, "17 Kernsätze zur Logik" (*c*. 1906). Contrast Ricketts ("Objectivity and Objecthood", 81): "There is, as far as Frege is concerned, nothing to be said about the justification for our recognition of those basic laws of logic to be truths"; and Weiner (*Frege in Perspective*, 71–2). Frege says a good bit about the epistemology of belief in the basic laws, scattered through his writings. I shall not discuss these passages in this paper, however.

342–343, "Der Gedanke", O 58–59): a) psychological explanation of belief or judgment, including an account of its acquisition, (b) justification of our belief or judgment, and (c) grounding for logical truth. Frege always lays aside psychological explanation. But he repeatedly discusses the justification of "our" belief or judgment in logical truths as well as the grounding of logical truth. Understanding grounding of truth is a matter of understanding the natural order of truths, which is independent of thinking or practice, and the "same for all men" (*FA* p. ii; *FA* §17; *BL* 13, 15; *GG*, pp. xvi, xvii). One understands the grounding of truth when one understands the natural order of logical and mathematical proofs, and the primitive truths on which such proofs rest. What grounds logical truths are the primitive logical truths.

One of Frege's primary motivations for understanding logical truth and the proof structure of logic was to understand the nature of justification for our mathematical judgments. In *Foundations of Arithmetic* Frege begins in §§1 and 2 by announcing an interest in the proof structure of mathematics. But he immediately associates this structure with the question of the justification of belief. In §2 he says that the aim of proof is, partly, to place a proposition beyond doubt. In §3 he says that "philosophical motives" underly his inquiry into the foundations of mathematics: The motives turn out to center on answering the question "Whence do we derive the justification of our assertion [of mathematical truths]"? The question of whether arithmetic is analytic turns out to concern the justification for making a judgment. He refines this to read, it concerns "the ultimate ground upon which rests the justification for holding [a proposition] to be true".

What is important about this passage is not only Frege's concentration on justification for judgment, but also his belief that the justification of an arithmetical judgment derives from the mathematical foundation (*Grund*)—from the primitive truths. The problem [of finding the justification for assertion or judgment], he says, is to be solved by "finding the proof of the proposition and following it back to the primitive truths" (*FA* p. 4). One might ask, how can a problem of the justification of our beliefs or judgments be solved by citing primitive truths? How can such truths be primary in an account of justification?

Frege's line is made clearer in "The Thought" where he characterizes laws of truth as general laws which concern not "what happens" but "what is". Speaking of these laws about "what is" in the third realm, Frege says that "from the laws of truth there *follow* ["ergibt sich"—a non-technical term] prescriptions about asserting, thinking, judging, inferring" (*CP* 351; *KS* 342, "Der Gedanke", O 58). Frege then calls these prescriptive epistemic laws "laws of thought" (*CP* 351; *KS* 342, "Der Gedanke", O 58). This is a paradigmatic Platonic direction of explanation: from what *is* in an abstract realm to what is reasonable. What could be the nature of this derivation from general laws of truth—which concern logical objects and functions—to prescriptive laws about judgment? Frege writes: "[the laws of truth] are boundary stones set in an eternal foundation, which our thought can overflow, but never displace. It is because of this that they have authority for our thought *if it would attain to truth*" (*BL* 13; *GG* p. xvi).

Is it contingent that a judging subject "would attain to truth"? Frege is certainly insistent that the laws of truth are independent of their being taken to be true by anyone (*BL* 13, 15; *GG*, pp. xvi, xvii). Moreover, he thinks it not contradictory to suppose something's being true which everyone takes to be false (*BL* 13; *GG* pp. xv–xvi). On the other hand, Frege sees judgment as an advance from thought content to truth-value. The function or aim of judgment is to reach truth. So to be a judging subject, one must have this aim or function insofar as one makes judgments. In this sense, the prescriptions of the laws of truth must apply to the judgments of judging subjects.

There is a second way in which Frege thinks that there is a deep, non-contingent relation between the laws of truth and prescriptive laws about judgment. To be rational, he thinks, one must be disposed to acknowledge the simplest logical truths. Judgments in contradiction with the laws of logic would constitute a kind of madness (*BL* 14; *GG*, p. xvi). In fact, Frege appears to believe that failure to acknowledge primitive logical laws, like the principle that every object is identical with itself, and even certain truths of arithmetic, would throw thought into confusion and undermine the possibility of genuine judgment and thought (*FA* p. 21). This suggests that he was inclined to believe that a disposition to acknowledge basic logical truths and inferences—and a disposition not to deny non-basic but relatively simple truths of arithmetic—is a condition not only for being rational but for being a judge or thinker at all.

Here it is worth looking very carefully at a famous passage in *Basic Laws*. Frege considers a supposed possibility in which beings have laws of thought (prescriptions for judgment) that contradicted ours. He claims that such beings would exhibit a "hitherto unknown type of madness", and indicates that such beings' procedures for taking things to be true would be in radical disaccord with the laws of truth (*BL* 14; *GG* p. xvii). Shortly afterwards, he writes:

> If we step away from logic, we may say: we are compelled to make judgments by our own nature and by external circumstances; and if we do so, we cannot reject this law—of Identity, for example; we must acknowledge it unless we wish to reduce our thought to confusion and finally do without all judgment whatever. I shall neither dispute nor support this view; I shall merely remark that what we have here is not a logical consequence. What is given is not a ground (*Grund*) for something's being true, but for our taking it to be true. Not only that: this impossibility of our rejecting the law in question hinders us not at all in supposing beings who do reject it; where it hinders us is in supposing that these beings are right in so doing . . . Anyone who has once acknowledged a law of truth has by the same token acknowledged a law that prescribes the way in which one ought to judge, no matter where, or when, or by whom the judgment is made. (*BL* 15; *GG* p. xvii)

Frege is taking a hands-off attitude toward the epistemological issues for the purpose of his mathematical treatise. But, given his own beliefs, what does he neither dispute nor support in the view he states? Some have thought that in citing the limits of logic, he is prescinding from any judgment about grounds for our taking something to be true, as opposed to the ground for its being true. Some have even held that grounds for our taking something to be true are thought by Frege to be psychologistic, and of no interest to him. These are serious misreadings of the passage.

Understanding grounds for our *taking* something to be true had long been what motivated his inquiry into the foundations of arithmetic. (Cf. especially *FA* p. 3, where he uses exactly the same German terms as he does in the above-cited passage: grounds for taking [holding] something to be true.) One page earlier in *Basic Laws* Frege characterized the laws of logic in the double way I have described: not only as laws of truth but as laws that "prescribe the way in which one ought to think" (*BL* 14; *GG* p. xvi)—as laws of thought. What Frege takes no position on is whether we are compelled to acknowledge the laws by our own nature and by external circumstances.

This is indeed a psychological matter. He thinks that any such psychological law would admit of conceivable exceptions—mad beings that do reject the law. But where he writes, "we must acknowledge it unless we wish to reduce our thought to confusion and finally do without all judgment whatever", he is speaking in his own voice. For he had already indicated that he believes that renouncing the laws of arithmetic (which are less basic for him than the basic laws of logic) would be to reduce thought to confusion and make thinking impossible (*FA* p. 21). Frege thinks that acknowledging these laws, at least implicitly in one's actual thinking, is necessary for having reason and for being a non-degenerate thinking and judging subject. (He apparently believed that although a mad person could reject a law, abiding by such rejection would reduce thought to confusion, and by degrees undermine judgment altogether.) These are normative not psychological judgments. Although they are not logical consequences, they are part of Frege's epistemic view.

So let us summarize the view that Frege maintains. He holds that justification for holding logical laws to be true rests on and follows from primitive laws of truth. He spells out this dependence of epistemic justification on the laws of truth in two ways. He thinks that laws of truth indicate how one ought to think "if one would attain to truth". But a judging subject necessarily would attain to truth, insofar as it engages in judgment. So any judgment by a particular person necessarily is subject to the prescriptive laws set out by the primitive laws of logic. One is justified in acknowledging them because doing so is necessary to fulfilling one's aim and function as a judging subject.

Frege's second way is: acknowledgment of certain laws of truth is necessary for having reason and for engaging in non-degenerate thinking and judging. One is rationally entitled to judge the primitive laws of logic to be true because the nature of reason—and even non-degenerate judgment—is partly constituted by the prescription that one acknowledge at least the simple and basic laws of truth. To put it crudely, reason and judgment—indeed mind—are partly defined in terms of acknowledging the basic laws of truth.[17]

Our problem was to explain how, for Frege, mere reason could give grounds to believe that a subject matter is any particular way, given that the subject matter is atemporal, causally inert, and independent of thinking about it. Most current approaches to the substantive problem look for some analog of causal interrelation in our mathematical knowledge. More traditional views—both Platonic and idealist—see the relation as individuative or constitutive.

An idealist line is to make the subject matter constitutively dependent on thinking, synthesis, or practice. Frege's line is to hold that, although the laws of truth are independent of judging subjects, judging subjects are in two ways not independent of the laws of truth. First, to be a judging subject is to be subject to the prescriptions of reason, which in turn are provided by the laws of truth (logic). For judgment has the function of attaining truth; and the laws of logic—which are constituted by atemporal thoughts and atemporal subject matter—provide universal prescriptions of how one ought to think, given that one's thinking has the function of attaining truth. Second, being a judging subject is to have or have had some degree of reason. Having or having had some degree of reason requires acknowledging, at least implicitly in one's thinking, the simplest, most basic logical truths and inferences; and doing so commits one to an atemporal subject matter. Questions of "access" to the third realm are on reflection seen to be misconceived. For, to reverse somewhat Gertrude Stein's dictum about Oakland, there is no there there.

Why was this line not more prominent in Frege's philosophy? He thought that his primary contribution lay in identifying primitive truths and inference rules, and in deriving arithmetic from them. He accepted the traditional rationalist–Platonist line about the relation between reason and primitive truths. He did not think it needed substantial elaboration. Like Frege, I think that this neglected line is not to be dismissed. Unlike Frege, I think it may be worth developing.

[17] One can see this view alluded to in the passage where Frege claims that logic can, with anti-idealist and anti-psychologistic qualifications, be seen as the study of not minds but Mind (*Geist*) (*CP* 369; *KS* 359, "Der Gedanke", O 74). One can also see it in his claim: "We might with alteration of a well-known proposition say: the proper object of reason is reason. In arithmetic we are not concerned with objects which we come to know as something alien from without through the medium of the senses, but with objects given directly to reason, which as her most proper objects are completely transparent to her" (*FA* p. 115). Cf. also *B* § 23 and *FA* §26. These quotes are not idealist, as they have sometimes been taken. They are expressions of the view that the basic forms and objects of logic constitutively inform minds, and help define what it is to be mind or reason.

9 *Frege on Knowing the Foundation (1998)*

From the start of his career Frege motivated his logicism epistemologically. He saw arithmetical judgments as resting on a foundation of logical principles, and he saw the discovery of this foundation as a discovery of the nature and structure of the *justification* of arithmetical truths and judgments. Frege provides no focused and sustained account of the foundation, or of our epistemic relation to it. It is clear from numerous remarks, however, that he saw the foundation as consisting of primitive logical truths, which may be used as axioms and which are self-evident. He thought that they are in need of no justification from any other principles. Logical and arithmetical principles other than the self-evident ones are justified by being provable from self-evident axioms together with self-evident definitions and self-evident rules of inference.

At first glance, Frege may seem to be a prototypical representative of the Euclidean, rationalist tradition in the epistemology of logic and mathematics. But further scrutiny reveals a more complex situation. Although Frege regards the axioms as self-evident, he expressed a sophisticated modern awareness of the fact that what can seem obvious may turn out not even to be true. (The distinction between self-evidence and obviousness will be discussed.) Ironically, this awareness led to at least dim anticipations of the problem with his fifth axiom, a problem that decimated his logicist project. Frege was aware that principles that he put forward as axiomatic—even some that, unlike Axiom V, have endured as basic principles of logic—were not found to be obvious by his peers. In arguing for his logic he made use of methods that were explicitly pragmatic and contextualist, in senses that will

The following abstract appeared in the original publication:

The paper scrutinizes Frege's Euclideanism—his view of arithmetic and geometry as resting on a small number of self-evident axioms from which non-self-evident theorems can be proved. Frege's notions of self-evidence and axiom are discussed in some detail. Elements in Frege's position that are in apparent tension with his Euclideanism are considered—his introduction of axioms in *The Basic Laws of Arithmetic* through argument, his fallibilism about mathematical understanding, and his view that understanding is closely associated with inferential abilities. The resolution of the tensions indicates that Frege maintained a sophisticated and challenging form of rationalism, one relevant to current epistemology and parts of the philosophy of mathematics.

be discussed. The relations between these aspects of Frege's position and his views on self-evidence are worth understanding.

Moreover, whereas Frege maintained with the rationalist tradition that understanding certain truths justifies one in believing them, he developed an original conception of what is necessary for understanding. In particular, understanding presupposes inferential abilities, even where those inferences are not needed for epistemic support of belief in the thought that is understood.

The purpose of this essay is to investigate Frege's conception of the knowledge and justification of the primitive logical truths, from which he intended to take his axioms. Frege did maintain the primary tenets of traditional rationalism. But the nature of his work and his place in history gave his philosophical genius materials for supplementing the traditional view and developing neglected aspects of it, in such a way as to make it less vulnerable to some of the traditional objections. In fact, I think that the view that he developed is of current philosophical interest.

I shall begin in §I by outlining the basics of Frege's rationalist position—his Euclideanism about his axioms. In §II, I discuss how he introduces his axioms in his two main works of logic. I discuss puzzles about the relation between these modes of introduction and his belief that the axioms are self-evident. In §III, I expound "pragmatic" elements in his work that do not seem to fit with the traditional view. §IV is devoted to a further development of Frege's rationalism, with special attention to the notion of self-evidence and to how the Euclidean and "pragmatic" tendencies relate to one another.

I

Frege opens *Begriffsschrift* (1879) with the following statement

> In apprehending a scientific truth we pass, as a rule, through various degrees of certitude. Perhaps first conjectured on the basis of an insufficient number of particular cases a general proposition comes to be more and more securely established by being connected with other truths through chains of inferences, whether consequences are derived from it that are confirmed in some other way or whether, conversely, it is seen to be a consequence of propositions already established. Hence we can inquire, on the one hand, how we have gradually arrived at a given proposition and, on the other, how it is finally to be most securely grounded [or founded: *begrundet*]. The first question may have to be answered differently for different persons; the second is more definite, and the answer to it is connected with the inner nature of the proposition considered. The most reliable way of carrying out a proof, obviously, is to follow pure logic, a way that disregarding the particular characteristics of objects depends solely on those laws upon which all knowledge rests. Accordingly, we divide all truths that require grounding (*Begrundung*) into two kinds, those for which the proof can be carried out purely by means of logic and those for which it must

be supported by facts of experience. But that a proposition is of the first kind is surely compatible with the fact that it could nevertheless not have come to consciousness in a human mind without any activity of the senses. Hence it is not the psychological genesis but the best method of proof that is at the basis of the classification. (*B* Preface)[1]

This passage contains a number of recurrent themes. Frege draws a sharp distinction between psychological genesis and grounding. Grounding is firmly associated for Frege with justification—epistemic warrant or "support". "*Grund*" in German means not only ground, but also reason. So a grounding or founding is naturally associated with reasons. Frege sees proof as the primary relevant form of justification involved in grounding or founding a truth. The nature of the justification of a proposition is "connected with the inner nature of the proposition". The most reliable justification is one that "can be carried out purely by means of logic", one that rests solely on laws of logic "upon which all knowledge rests".

According to Frege's implicit conception of justification, justification or foundation is associated with propositions, eventually in Frege's mature work with thoughts (thought contents), not—or at least not explicitly—with beliefs of individuals.[2] The justification for a proposition consists in the best method of proving it—in an actual abstract proof structure which constitutes the *possibility* of carrying out a certain sort of proof.

Frege regards a justification as a structure which may be understood or psychologically mastered in different ways by different people—or perhaps not understood at all. The structure is associated with the "inner nature of the proposition", not with individual abilities or states of thinkers.[3]

As Frege implies, not all propositions require a grounding. Some stand on their own. In *B* §13, Frege continues:

[1] References to corresponding passages in English and German are separated by a slash mark. The following works of Frege will be cited with these abbreviations: *The Basic Laws of Arithmetic* (*BL*); *Begriffsschrift* (*B*); *Collected Papers on Mathematics, Logic, and Philosophy* (*CP*); *The Foundations of Arithmetic* (*FA*); *Die Grundgesetze der Arithmetik* (*GG*); *Kleine Schriften* (*KS*); *Nachgelassene Schriften* (*NS*); *Philosophical and Mathematical Correspondence* (*PMC*); *Posthumous Writings* (*PW*); *Translations from the Philosophical Writings of Gottlob Frege* (G & B); *Wissenschaftlicher Briefwechsel* (*WB*). "O" indicates pagination in the original publications of Frege's articles.

[2] Later—in *B* (§13)—he associates it with "judgments of pure thought", which are seen as abstract, possible acts of judgment associated with ideal logical thinking.

It is unclear how Frege regarded inductive arguments. Sometimes (*FA* §2) he contrasts inductive arguments with proofs. In other places, he seems to see them as ultimately deductive arguments: as starting with singular statements about particulars, together with a general law of induction, and yielding empirical laws as conclusions. He does not say how the conclusions of such "proofs" should be seen. That is, he does not indicate whether the laws are less than conclusively established, or whether the non-conclusive nature of the argument is built into the statement of the conclusion in some way. (Cf. *FA* §3.)

[3] For another early association of logical principles with justification, see *NS* 4, 6/*PW* 4, 5, "Logik" (zwischen 1879 und 1891).

It seems natural to derive the more complex of these judgments from simpler ones, not in order to make them more certain, which would be unnecessary in most cases, but in order to let the relations of the judgments to one another emerge. Merely to know the laws is obviously not the same as to know them together with the connections that some have to others. In this way we arrive at a small number of laws in which, if we add those contained in the rules, the content of all the laws is included, albeit in an undeveloped state. And that the deductive mode of presentation makes us acquainted with that core is another of its advantages. Since in view of the boundless multitude of laws that can be enunciated we cannot list them all, we cannot achieve completeness except by searching out those that, *by their power*, contain all of them. Now it must be admitted, certainly, that the way followed here is not the only one in which the reduction can be done. That is why not all relations between the laws of thought are elucidated by means of the present mode of presentation. There is perhaps another set of judgments from which, when those contained in the rules are added, all laws of thought could likewise be deduced.

Frege here elaborates his view of a justificational structure. He sees simpler judgments as tending to be more basic than complex ones (*NS* 6, 36/*PW* 6, "Logik" (zwischen 1879 und 1891); *PW* 36, "Booles rechnende Logik und die Begriffsschrift" (1880–1881)); *KS* 165/*CP* 180, "Rezension von: Georg Cantor, Zur Lehre vom Transfiniten", O 271–272). More complex rules of inference are similarly to be resolved into simple, basic ones (*FA* §90; cf, *GG* pp. vi–vii/*BL* 2–3).

There is an analogous structure among parts of thoughts that would undergo definition. Frege writes:

In the case of any definition whatever we must presuppose as known something by means of which we explain what we want understood by this name or sign. . . . To be sure, that on which we base our definitions may itself have been defined previously; however, when we retrace our steps further, we shall always come upon something which, being a simple is indefinable, and must be admitted to be incapable of further analysis. And the properties belonging to these ultimate building blocks of a discipline contain, as it were in a nutshell, its whole contents. (*KS* 104/*CP* 113, "Über formale Theorien der Arithmetik", O 96; cf. also *KS* 289–90/*CP* 302, "Über die Grundlagen der Geometrie ii" (1903), O 303)

Both of the last two quoted passages show that Frege sees the structure as residing in the nature of the laws or contents. Basic, foundational laws have—by their content, or nature—the power of proving or entailing the others. As we shall see shortly, Frege thought that such laws are *unprovable*. Basic thought components are those that by their content, or nature, have the power of defining others. Such "simples" are *indefinable*.

The *Begriffsschrift* § 13 passage indicates that Frege thinks that there are other laws than those he *takes* as basic *for the purpose of formalizing a system* of logic, which also have the power of yielding all the other laws as logical consequences. These laws also are basic *in themselves*. They do not need proof. Thus there appear to be more "basic" laws in logical reality than are needed to provide a simple and elegant formal presentation of logic (cf. *KS* 391/*CP* 404, "Gedankengefüge", O 48–49; *NS* 221–2/*PW* 205, "Logik in der Mathematik" (1914)). Studying these alternative ways of deriving the non-basic theorems would yield deeper insight into the various relations among propositions.

Frege says little more about basic laws in *Begriffschrift*. His concern is to provide the logical tools to enable one to determine the exact nature of proof or justification. More particularly, he is concerned that failure to formalize the steps in a proof tends to lead to one's making inferences that inarticulately amalgamate several steps into one.[4]

The reason that this tendency is a source of concern for him is that it may lead one to think that the inference depends on "intuition", on a non-logical cognitive capacity, when in fact it is purely logical. The logical character of the inference may be obscured by the fact that too many logical steps are lumped together, without a clear account of any of them. The result would be a tendency to package ill-understood inferences under the ill-understood category "intuition".[5] A tendency to count logical inferences non-logical is, of course, an obstacle to Frege's *logicist thesis*: the thesis that arithmetic can be reduced to logic together with definitions of arithmetical terms in logical terms.

Frege's concern throughout is with understanding the nature of the justification or grounding of arithmetic. Both at the beginning and at the end of the Preface to *Begriffschrift* Frege announces his logicist aims. In both places, he associates understanding logical and arithmetical structures with understanding the justification or ground of logical and arithmetical propositions. Understanding proof structure was for Frege equivalent to understanding the structure of justification (cf. also *FA* § 3). Frege's project was a project in the theory of knowledge.

The same picture of the justification of arithmetic and logic as resting on a small group of logical laws appears again in *The Foundations of Arithmetic* (1884). In § 1, Frege commends "the old Euclidean standards of rigour". He writes, "The aim of proof is . . . not merely to place the truth of a proposition beyond all doubt, but also to afford us insight into the dependence of truths upon one another" (§ 2). This is equivalent to affording us insight into the nature of justification: "After we have convinced ourselves that

[4] Cf. Preface of *B* and §23. Cf. also *FA* §§1–3; "On Mr. Peano's Conceptual Notation and My Own" (1897), *KS* 221 ff./*CP* 235 ff., O 362 ff.

[5] The notion of intuition carries the special technical understanding of the term associated with Kant. Pure intuition (roughly the spatio-temporal structure of perceptual capacities) was supposed by Kant to be a source of apriori knowledge that is not logical. Frege agreed with this view as applied to geometry. But he thought appeal to it obscured the logical foundation underlying arithmetic. No doubt Frege believed that appeals to a non-technical notion of mathematical intuition (insight, sense of the obvious) also obscured the logical character of arithmetic axioms and proof.

a boulder is immovable, by trying unsuccessfully to move it, there remains the further question, what is it that supports it so securely?". Again, Frege invokes the task of finding the primitive truths to which everything else can be reduced. In § 3, Frege indicates that whether arithmetic is "analytic"—whether it can be reduced to logic together with definitions—is a question about.

> the ultimate ground upon which rests the justification (*Berechtigung*) for holding [a proposition of arithmetic] true.... The problem becomes, in fact, that of finding the proof of the proposition, and of following it up right back to the primitive truths. If, in carrying out this process, we come only on general logical laws and on definitions, then the truth is an analytic one.... (*FA* §64; also *KS* 104/*CP* 113–14, "Über formale Theorien der Arithmetik", O 96; *KS* 289–90/CP 302, "Über die Grundlagen der Geometrie" (1906), O 303)

In discussing the crucial definitions later in the book, he manifests clear concern that the definitions respect conceptual priority—which is ultimately justificatory priority.

Frege offers no large systematic discussion of knowledge of the foundations, of the primitive truths on which the justificational structure rests. In fact, he neglects to formulate his notions of analyticity and apriority so as to either include or rule out the foundations of logic.[6] But as we shall see, Frege makes many remarks about knowledge of the foundations, which enable one to put together an account of his view.

Frege assumes that the foundations are "general" propositions—not propositions about particulars (*FA* §§3, 5). He thinks primitive truths of logic must not have concepts that are peculiar or special to any non-universal subject matter (as the concepts of geometry do—in their relevance to space). And he thinks that one of the aims of his project is, by showing how the truths of arithmetic rest on primitive truths, to place the truth of the propositions of arithmetic beyond all doubt—on the firmest possible foundation (*FA* §2). There is much more in *Foundations* that gives one insight into Frege's view of the epistemic status of the axioms. But I want to postpone discussion of some of this material until §IV.

In the two decades after *Foundations* there are further remarks that indicate Frege's Euclidean views about the basic truths. In 1897 he writes:

> I became aware of the need for a conceptual notation when I was looking for the fundamental principles or axioms upon which the whole of mathematics rests. Only after this question is answered can it be hoped to trace successfully the *springs of knowledge* [my emphasis] upon which this science thrives. Even if this question belongs largely to philosophy, it must still be regarded as mathematical. The

[6] Michael Dummett *Frege: Philosophy of Mathematics* (Cambridge, Mass.: Harvard University Press. (1991), ch. 3) notices this. I think Dummett is right that this is an oversight on Frege's part—or at any rate that Frege intended no epistemic slight to the foundations. For he repeatedly connects foundations to justification, and repeatedly calls the foundations self-evident, as I shall discuss below.

question is an old one: apparently it was already being asked by Euclid. (*KS* 221/*CP* 235, "Über die Begriffsschrift des Herrn Peano und meine Eigene" (1896), O 362)

Frege requires that for something to count as an axiom, it must be true, certain, and unprovable: "Traditionally, what is called an axiom is a thought whose truth is certain without, however, being provable by a chain of logical inferences. The laws of logic, too, are of this nature" (*KS* 262/*CP* 273, "Über die Grundlagen der Geometrie i", (1903), O 319). Frege sometimes counts these three features—truth, certainty, and unprovability—as constituting the Euclidean *meaning* of "axiom" (*KS* 283/*CP* 295, "Über die Grundlagen der Geometrie ii" (1903), O 296; cf. *KS* 313/*CP* 328, "Über die Grundlagen der Geometrie ii" (1903), O 398; *NS* 183/*PW* 168, "Über die Euklidische Geometrie", (1899–1906?); *NS*/266–7/*PW* 247, "Logik in der Mathematik" (1914)).

This conception is indeed traditional. But it has been obscured by subsequent developments. Many use the term "axiom" now in a way that would allow for no incoherence or inappropriateness in talking of false axioms. On this usage, axioms are basic principles of a theory, something proposed by human beings and capable of being found to be mistaken. We naturally speak of Frege's Axiom V, which turned out to be inconsistent, but certainly was an "axiom" (in our less traditional sense) of his system of logic.

Frege's Euclidean conception of axioms takes them to be true first principles—basic truths—which might or might not be discovered or proposed by human beings. Something's being an axiom is not primarily a matter of being part of a theory. Frege does think that whether a basic truth is an axiom depends on its being used as an axiom, as starting point in a system of derivation. (This differentiates the notion *axiom* from the notion *basic truth*.) But for Frege a necessary condition on something's being an axiom is its being a basic truth—in particular, a foundational part of a mathematical or logical structure which it is the purpose of logicians and mathematicians to discern and express. If a proposition is not *true*, then it cannot possibly stand at the foundation of these structures. So it cannot be an axiom. Being an axiom, on Frege's conception, is necessarily being part of the foundation of these structures.

The sense in which Frege thought that the primitive truths are certain is more complex. The notion of certainty figures primarily in Frege's generalized motivational sections. He thought that basic truths' certainty grounds the certainty of theorems derived from them. It is clear that he did not regard this certainty in a purely psychological sense. He did not think that just anyone who had thought about axioms had to be maximally confident that they are true. The reason this is clear is that he knew that many mathematicians who had thought about the axioms of Euclidean geometry did not think of them (or in some cases, any other mathematical propositions) as true. He was also aware that many logicians did not accept his system of logic (*KS* 262 ff./*CP*

273 ff., "Über die Grundlagen der Geometrie i" (1903), O 319 ff.; *NS* 183–4/ *PW* 168–9, "Über die Euklidische Geometrie" (1899–1906?)).[7] He seems to talk of certainty as a property of the logical and mathematical truths. For example, in *Foundations* he associates certainty with the immovability of a boulder, and with being beyond reasonable doubt (*FA* §2). There is some reason to think that Frege held that empirical propositions lack this feature. (Compare *KS* 115/*CP* 125, "Über das Trägheitsgesetz", O 148, with *NS* 286–8/*PW* 267–9, "Erkenntnisquellen der Mathematik und der mathematischen Naturwissenschaften" (1924/1925)).[8]

Frege regarded not only axioms but most of the *theorems* of logic and at least the more ordinary theorems of arithmetic as being certain (*B* § 13; *FA* § 2). He had almost no sceptical impulse, and he seemed to have regarded quite a lot of non-basic truths—as well as basic ones—as certain. Certainty appears to mean something like *beyond a reasonable doubt by someone who fully understands the relevant propositions*.

In addition to requiring that axioms be true and certain, Frege required that axioms be unprovable. Now this term "unprovable" seems to have two different but compatible interpretations for Frege. In his mature work Frege is very explicit that axioms are abstract thoughts—thought contents—not sentences. What are proved are, similarly, thoughts (*Gedanken*) not sentences. But on one interpretation of "unprovable", axioms are unprovable relative to a system—an ordering marked out by human beings—of true, certain thoughts. That is, they are starting points in a certain system of derivation. We noted earlier that a thought is an axiom only if it is used as an axiom. Only true and certain thoughts can be axioms in Frege's sense of "axiom". Truth and certainty are not relative to a system. But on this interpretation of "unprovable", unprovability is relative to a system. Whether a truth is unprovable depends partly on whether it is taken to be an axiom in a given logical theory. (One-step "proofs" are not allowed.) In one system a thought may be an axiom, and in another system the same thought may be a theorem. The point is at least suggested in *Begriffsschrift* §13. But Frege makes this interpretation of "unprovable" fully explicit only late in his writings (*NS* 221–2/ *P W* 205, "Logik in der Mathematik" (1914)). It is, however, clearly available to him throughout, and he may have had it in mind in other passages. In this sense, unprovability is not an intrinsic feature of a truth, but is rather a status accorded to it by virtue of its place in a given logical system. The same truth may have another place (as theorem) in another system.

I think, however, that this interpretation cannot provide all that Frege meant by "unprovable". In any case, being true, certain, and unprovable in

[7] Frege's awareness of doubts about his logical system dates from the reception of his *Begriffsschrift*. Cf. his articles against the Booleans, *NS* 9–59/*PW* 9–52. It also appears in the Introduction to *BL*. Cf. esp. *GG* p. xii/*BL* 9.

[8] Some passages in "Thoughts" (1918) (*KS* 342–61/*CP* 351–352, "Der Gedanke", O 58–59) suggest that he thought that the existence of the physical world was certain.

the sense just adumbrated is not sufficient for being an axiom. Those notions cannot jointly constitute the Euclidean meaning of "axiom" for Frege—as he says truth, certainty and "unprovability" do. (Cf. again (*KS* 283/*CP* 295, "Über die Grundlagen der Geometrie ii" (1903), O 296; cf. *KS* 313/*CP* 328, "Über die Grundlagen der Geometrie ii" (1903), O 398; *NS* 183/*PW* 168, "Über die Euklidische Geometrie" (1899–1906?); *NS* 266–7/*PW* 247, "Logik in der Mathematik" (1914).) In some passages where Frege calls axioms "unprovable", he clearly assumes that they are "*basic laws*" (e.g. *KS* 262/*CP* 273, "Über die Grundlagen der Geometrie i" (1903), O 319). Indeed, the connection between being an axiom and being a basic law is constant throughout Frege's work. Not just any true and certain thought would be a basic law, and appropriately taken as an axiom, merely by virtue of being *taken* as a contingently "unprovable" starting point for a system of proof. Theorems of arithmetic were regarded by Frege as true and certain, and they could be arbitrarily taken as starting points in a system of arithmetic. But Frege would not regard them as axioms.

In *Foundations* §5 Frege argues against those who regarded numerical equalities about particular numbers as unprovable. The argument takes the issue to be about a matter of fact, not one of choice. Even if his opponents took some or all such equalities as starting points for their systems of derivation, Frege would not concede that they are "unprovable" in the sense that he is using the term. He argues that all such mathematical propositions are provable, and this argument is meant not just to advertise his own intention not to take them as axioms, but to indicate that logicism is true about the nature of arithmetic: they are provable from basic laws of logic. No arithmetical statement is an axiom, because all are provable. Frege was seeking basic laws in a sense that transcends sociological facts about what (true, certain) propositions were used as starting points in actual systems of derivation.

The assumption that axioms are basic laws or basic truths can be explicated in terms of Frege's characterization of primitive general laws as being "neither capable nor in need of proof" (*FA* §3). This phrase comes directly from Leibniz, from whom Frege probably got it (Leibniz, *New Essays on Human Understanding*, IV. ix. 2). The passage suggests a second interpretation of "unprovable". Leibniz certainly thought of basic truths as unprovable in a sense that goes beyond their being *taken* as starting points in a system of derivation. Leibniz thought them unprovable in the sense that they could not be justified by being derived from epistemically prior truths. In the relevant passage (*FA* §3), Frege like Leibniz is discussing a natural order of justification among truths. Frege thought that basic truths have to stand on their own in a natural structure of proof (justification, grounding). Again, he thought of this structure as independent of theories put forward by human beings.

Frege sees proof not fundamentally as just any structure of logical derivation, but as a form of justification or grounding. From this perspective, basic truths are unprovable in the sense that they cannot be grounded or given a justification by being derived from other truths. They can be derived, according to logical rules, from other truths within certain systems. But the derivations would not be justifications, groundings, or proofs in this epistemically fundamental sense. So from this perspective, basic truths are unprovable in the sense that they cannot be grounded or justified by being proved from other truths.

Moreover, basic truths do not need proof; they do not need justification or grounding through derivation from other truths. This epistemic feature of basicness is also fundamental to being an axiom. Frege writes, "it is part of the concept of an axiom that it can be recognized as true independently of other truths" (*NS* 183/*PW* 168, "Über die Euklidische Geometrie" (1899–1906?)). Axioms have to be basic truths. To be basic, a truth must not need or admit of proof. To be basic, a truth cannot be justified by being derived from epistemically more fundamental truths, and yet does not need such justification because it can be recognized as true independent of it.

Basic truths are, of course, certain. But Frege's notion of axiom still does not collapse into his notion of basic truth. As is evident from *Begriffsschrift* §13, quoted above, Frege thought that different principles from the ones he proposed could provide an adequate basis for a formal logical system and could suffice to derive all the theorems of logic or arithmetic. In the context of the passage and his later work, Frege is obviously conceiving of a natural epistemic order of derivation. So he must have thought that there are more basic logical truths than are needed to derive and justify all the (non-basic) truths of logic and arithmetic. The epistemically basic truths overdetermine the whole system of logical-arithmetic truths. Thus although some basic truths might be expressed as theorems in a formal system, they are not, from the point of view of the natural order of justification or proof, essentially derivative. They are essentially basic. But in the relevant system, they would not be axioms. Thus not all basic truths that are candidates for being axioms are, relative to a given system, in fact axioms. They would not be "unprovable" in the first of the two senses that we distinguished, even though they are "unprovable" in the second sense.

I conjecture that according to Frege's fully elaborated notion of axiom, an axiom is a thought that is true, certain, basic, and unprovable in the first of our senses. Or equivalently, an axiom is a thought that is true, certain, unprovable in both of our senses, and not in epistemic need of proof.

As I have noted, Frege's point that axioms (and basic truths) do not need proof is associated with, and probably equivalent to, his claim (properly understood) that an axiom—and a basic truth—"can be recognized as true independently of other truths" (*NS* 183/*PW* 168, "Über die Euklidische Geometrie" (1899–1906?)). He clearly thinks that the axioms of geometry and the axioms of logic have this feature. And at least at the end of his career—probably throughout it—he thought that sense experience statements about the physical world lack it (*NS* 286–8/*PW* 267–9, "Erkenntnisquellen der Mathematik und der mathematischen Naturwissenschaften" (1924/1925)). The requirement that axioms do not need proof is closely related to his requirement that axioms be *self-evident* (*FA* §§5, 90; *GG* ii. §60; G&B 164 ("*selbst-verständlich*"); *GG* 253/*BL* 127 ("*einleuchtend*")). Indeed, self-evidence must partly be understood in terms of recognizability as true independently of recognition of other truths. Sufficient evidence to make believing them rational is carried in these individual truths themselves.

The meaning of the various modal notions that Frege uses in the phrases "unprovable" (in our second sense) and "can be recognized . . . independently" in his requirement that axioms and basic truths have a sort of self-justification is complex. Understanding the modal notions is closely related to understanding Frege's notion of self-evidence. Developing a deeper understanding of this latter notion will occupy us in §IV. For now, it is enough to see the general shape of Frege's rationalism.

Justification resides in a proof structure that is independent of language and theory, but has an objectivity and reality that waits to be discovered. The proof structure involves basic truths which are justified in themselves, without need of proof. They contain simple indefinable concepts and are self-evident. Similarly, basic inference rules (*FA* 90) and definitions (WB 62/*PMC* 36, Frege to Hilbert, 12/27/1899; *KS* 263/*CP* 274, "Über die Grundlagen der Geometrie i" (1903), O 320; *KS* 289–90/*CP* 302, "Über die Grundlagen der Geometrie ii" (1903), O 303) are required to be self-evident. The certainty, or rational unassailability, of theorems which they entail is derivative from the certainty and self-evidence of the basic truths.

So far, we have a fairly familiar picture of Frege's indebtedness to the Euclidean rationalist tradition. I want to turn now to elements in Frege's views that are in apparent discord with his Euclidean epistemology. I will try to show wherein these further elements are compatible with and indeed enrich his rationalism.

II

Frege's practice in introducing his axioms in *The Basic Laws of Arithmetic* is at first glance at odds with his Euclidean commitments. Although Frege requires axioms to be unprovable, self-evident, and not in need of proof, he seems to provide *arguments* for at least some of them when he introduces them in *Basic Laws*.

For example, in §12 Frege introduces the material conditional as a function with two arguments, whose value is the False if the True be taken as the first argument and any object other than the True be taken as the second argument;

and whose value is the True in all other cases. Then in §18 he introduces Axiom I. He writes

By § 12,

$(\Gamma \to (\Delta \to \Gamma))$

could be the False only if both Γ and Δ were the True while Γ was not the True. This is impossible; therefore

$\vdash (a \to (b \to a))$. (*BL* § 18)

Here Frege argues from the way he introduced the relevant logical function to the truth of the relevant axiom. So after introducing the material conditional as a function in terms of its truth table, he argues from the truth table to the truth of the axiom. In every case, except that of Axiom V (whose introduction is non-standard because it is defective), he appeals to versions of recognizably standard reasoning from truth conditions to the axioms.[9]

Now this practice may seem odd. It raises at least two puzzles. Frege seems to be *arguing* for axioms, which are supposed to be unprovable, self-evident, and not in need of proof. And he seems to be arguing for the axioms from *semantic claims*, whereas what purport to be the basic truths of logic are not about linguistic expressions or reference at all.

Let us begin with the first puzzle. It is clear that in a recognizable sense, Frege is giving arguments or demonstrations that are semantical—though we will have to qualify the sense. It seems equally clear that the arguments are arguments for the truth of the axioms. But it is less clear what the purposes of the arguments are. Moreover, they differ in interesting ways from one another.

In understanding the argument that Frege gives in §18 of *Basic Laws*, it is helpful to look at the counterpart passage in §14 of *Begriffsschrift*. There Frege writes,

$\vdash (a \to (b \to a))$

says "The case in which a is denied, b is affirmed, and a is affirmed is excluded". This is evident, since a cannot at the same time be denied and affirmed. We can also express the judgment in words thus, "If a proposition a holds, then it also holds in case an arbitrary proposition b holds."

This passage holds the key to the first puzzle. Frege clearly regards his argument as an elaboration of what is contained as evident in the axiom itself. It is an elaboration of an understanding of the thought, which is a basic truth. Frege does not see himself as starting with more basic truths—such as the

[9] I am grateful to Christopher Peacocke for drawing my attention to the interest of these arguments, and to Richard Heck for pointing out a mistake I had made about two of the individual axioms. Frege's lack of confidence in his official claim of self-evidence for Axiom V leads to a special treatment of its key primitive expression in §31. There he attempts to demonstrate that the course-of-values operator, when given an appropriate grammatical completion, always produces denotations. §31 raises numerous difficult interpretative questions that I pass over here.

principle of non-contradiction together with truths about the way the material conditional maps truth-values onto truth-values—and then justifying the axiom by reasoning to it from these resources. He sees himself as articulating in argument form what is contained in the very content of the basic truth he is arguing for. The truth is epistemically basic. Understanding it suffices for recognition of its truth.

Anyone who understands the truth can give the argument through understanding the material conditional. But the truth is not a conclusion of a proof, a structure of justification. For there is no justificational structure with truths more basic than the conclusion of the argument. The axiom is supposed to be unprovable in the second sense we have elucidated. Moreover, it does not need a proof. It is self-evident. It is evidence for its own truth. Epistemically, it is the truth's content, not the discursive argument, that is basic. The argument serves to articulate understanding of the thought content. It does so in a way that enables one to recognize that its truth is guaranteed by its content. The argument is not a derivation that justifies or grounds the thought in more basic truths.

In the *Begriffsschrift* passage, Frege indicates what the expression says, and he explains wherein what it says is evident. In the *Basic Laws* passage, he appeals to truth conditions associated with the relevant function—which he elsewhere identifies with the sense or content of the relevant proposition—and explains the truth of the axiom in the way he does in *Begriffsschrift*. The basic procedure is the same. Frege certainly saw the two cases as on a par. Frege is showing that the axiom's truth is evident from its content. The same point applies to all the introductory *arguments* for the axioms.

Let us broaden our perspective to include the second puzzle as well. In what sense does Frege argue from semantical principles in his introduction of Axiom I?

It should be noted that in the introduction of Axiom I (and of IV), Frege is not reasoning about symbols. So he does not take up a semantical perspective on his logical language at all. "Affirmed", in the *Begriffsschrift* formulation, converts in *Basic Laws* into "is the True", which is what the horizontal comes to express in Frege's mature work. The horizontal, or "is the True", is part of the expression of the axiom. "Is the true" is not a predicate of expressions or of axioms, truths, or thoughts. It denotes a concept of the truth-value the True—a function that takes only the True into the True. The predicate occurs in Frege's logic, along with the material conditional, negation, the universal quantifier, and so on. So Frege is not here using a meta-logical perspective in the modern sense.

The introduction of the material conditional in §12 of *Basic Laws* is, as I have noted, not a discussion of a symbol, but an explication of the material conditional as a function. And the argument introducing Axiom I in §18, which involves "is the True", also does not mention symbols at any point.[10] In that strict sense it is not semantical. Although the argument is rigorous, it is not epistemically fundamental. What is fundamental is the content of a single thought—the axiom. The argument simply articulates the self-evidence of the thought by expanding on what is involved in understanding it.

Two of the axioms (V and VI) contain *singular terms*—non-sentential terms formed from operations on function variables. The explications of the relevant singular terms in *Basic Laws* (§§, 9, 11) utilize "reference" or "denotation" (*Bedeutung*). These explications do mention expressions, and are unqualifiedly semantical. "Denotes" is a predicate that applies to linguistic expressions, and relates them to their denotations or referents. "Denotes" does not occur as a primitive logical term in any of the axioms.

Two other axioms (IIa and III) involve terms in places that we would normally reserve for not only variables but also singular terms in the sense explained in the previous paragraph. In these cases too, Frege's explication of the key expressions (the object-denoting term in universal instantiation and the identity predicate) invokes denotation. So again the explications are semantical.

In these four cases, Frege does not carry out any argument at all when he introduces the corresponding axioms (*Basic Laws*, §§18, 20). He simply cites the semantical explication of the expressions for the relevant functions (the universal quantifier, the identity sign, the definite description operator, and the course-of-values operator) and takes that appeal to make evident the truth of the corresponding axiom.[11] At most in introducing the axiom he reads through the axiom in its own terms, making sure that its key terms are understood.

So "denotes" occurs only in preliminary explications of logical symbols, not in any explicit argumentation that takes the axioms as conclusions. This may be significant. For there is, as noted, no argumentation for those axioms that we would nowadays read as containing singular terms. And *none* of the argumentation for the other axioms is semantical, in the sense that none of those arguments mention symbols. All of the arguments could be carried out within the language of the logic of *Begriffsschrift* (by avoiding the modal terminology). So *none* of the arguments that have the axioms as conclusions (as opposed to the preliminary explications of logical symbols or logical functions) are strictly semantical. They utilize only expressions (*modulo* the modal expressions) which could be formulated purely in Frege's logic. It is

[10] The modal expressions in the *BL* §18 passage need further investigation. I am inclined to think that they would be regarded by Frege, controversially of course, as inessential to arguments that he gives. Again, the comparison to the *B* passage is useful.

[11] It is true, of course, that Frege goes on to try to prove in §31 that singular terms involving the course-of-values operator have a reference. The proof is in some respects non-standard, and of course it fails. I take it that the attempt at a proof, which he would surely not have thought necessary for the other singular terms, was a sign of unease over the status and even truth of the proposed axioms involving the course-of-values operator.

not clear (or crucial to my purposes) what the significance of all this is. But I will return to these points.

I have noted that the explication of the material conditional is, in the first instance, an introduction of a function, not an explication of the logical symbol that denotes the function. That explication mentions no symbols. After that introduction, also in §12, Frege speaks of the denotation of the symbol for the material conditional. But it is the introduction of the function (without reference to expressions) that Frege cites in his argument.

The explications of negation (§6) and the definite description operator (§11) are like that of the material conditional. The function is introduced first, and the denotation of the sign is later noted in a meta-remark on the already established introduction of the function.

By contrast, the explication of identity (§7), which figures in the argument for axioms IV and III (§§18–19), does mention the symbol. And as I have noted, all the explications involving the singular terms (the free variables, terms formed from the definite description operator, and terms formed from the course-of-values operator) are full-blooded semantical explications.

This switching among methods shows, of course, that Frege was not *systematically* doing semantics of a language in the modern sense. He moves easily from using a denotation predicate that occurs outside his logical system to using a truth predicate that occurs in it. He moves easily from explications that introduce expressions to those that introduce the logical functions directly. In all cases, Frege regarded himself as both explaining his symbols and introducing functions that play the key role in the corresponding axioms.

Frege gives relatively rigorous, if not entirely systematic, semantical explications of his logical expressions. His explications systematically track semantical, if not necessarily model-theoretic, exposition. (I think Frege usually saw his language not as an uninterpreted symbolism, but as a perfect language carrying definite sense, differing from natural languages in that the sense and structure of the language serve rational inference ideally. For certain limited mathematical purposes, however, he does treat his language as a reinterpretable syntax in something like the model-theoretic fashion.) He expects such explications to aid in recognizing the truth of the axioms.

So Frege is doing several things at once. He is explaining the intended sense of his formulae by giving their truth conditions. He is implicitly justifying his logical language and the formulae that he uses by showing that they express logical truths and valid inferences, which are antecedently understood to be self-evident. And he is, in the arguments that derive the axioms, eliciting the self-evidence of the thoughts that are the axioms by bringing one to think through their content. But there is no sense in which he is justifying the axioms, the thoughts expressed by that language, through semantical argumentation. In fact, the argumentation he gives is within his logic (again, *modulo* the modal elements). Its function is not to justify the axioms by deriving them from prior truths, but to elicit understanding of them, an understanding that is supposed to suffice to enable one to recognize their self-evident truth.

Frege's introduction of his *methods of inference*—in contrast to the axioms—*is* systematically semantical. Using quotes, he introduces modus ponens this way:

> From the propositions [sentences, *Sätze*] "$\vdash (\Gamma \rightarrow \Delta)$" and "$\vdash \Gamma$" we may infer "$\vdash \Delta$"; for if Δ were not the True, then since Γ is the true $(\Gamma \rightarrow \Delta)$ would be the False. (*GG*/*BL* §14)

Frege is doing something different here from what he does when he gives the arguments for the axioms (which are thought contents—not formulae). What he calls a "method of inference" is a rule for moving from *sentences* (albeit fully meaningful sentences) to *sentences*. The argument that he gives is for the legitimacy of such rules. The argument is carried out within his logic (again assuming that the modal locutions are dispensable). But it is an argument about the legitimate use of his symbols.

Thus the arguments for the "methods of inference" differ from the "arguments" for the axioms. The arguments for the axioms in some cases rely on a preliminary explication of the logical notation. But they are not arguments that certain *expressions* (those that express the axioms) express logical truths. They are arguments whose conclusions are the axioms themselves, carried out in what we would call the object-language—using but not mentioning logical expressions. The axioms themselves are language-independent truths. The arguments for the methods of inference, by contrast, are arguments that certain transformations among meaningful sentences (not among what the sentences express) are truth-preserving. This contrast is important for understanding the epistemic function of Frege's argumentation for the axioms and methods of inference.

Later in *Basic Laws*, in criticizing formalist arithmetic, Frege implies a need to "justify" or "ground" "rules of inference", by appeal to the reference of the signs (§§90, 91, 94). Inevitably, the actual practice of proof must be formulated in terms of the permissible transitions among symbols expressing *Gedanken*. Frege is here writing of justifying or grounding methods of inference understood as ways of moving from symbols to symbols. He is justifying his introduction of his logical symbolism, not the language-independent logical principles or rules of inference that are expressed by the symbolism.[12]

[12] Richard Heck, "Frege and Semantics", in T. G. Ricketts (ed.), *The Cambridge Companion to Frege* (Cambridge: Cambridge University Press, forthcoming), and Jason Stanley, "Truth and Meta-Theory in Frege", *Pacific Philosophical Quarterly*, 77 (1996), 45–70), both cite these passages—§§90–4. The key thing to remember in reading these passages is that they concern rules about formulae, not rules concerning language-independent abstract thoughts. The passages oppose a formalist understanding of the *language* of arithmetic. In the context of §§90–4, Frege is claiming

Frege did not think of rules or methods of inference, in so far as they are transitions from symbols to symbols, as epistemically basic. It is not *these* methods of inference, rules about permissible transitions from expression to expression, that he refers to as self-evident when he discusses the epistemology of logic. For elsewhere Frege frequently indicates that rules of inference, in the strictest, most fundamental sense, have true thoughts, not expressions, as premises (*WB* 35/*PMC* 22, Frege to Dingler, 2/6/1917; *KS* 318–19/*CP* 334–5, "Über die Grundlagen der Geometrie iii" (1906), O 424). So logical inference is fundamentally a transition from abstract, language-independent thought contents (the eternal entities, *Gedanken*) to abstract thought contents. Rules of inference are fundamentally rules about such transitions. When Frege calls rules of inference "self-evident" (*FA* 90), he has in mind nothing about sentences, but logical principles of inference that the methods of inference—as principles about symbols—express. He assumes that modus ponens, understood as a method for moving from thoughts or judgments to thoughts or judgments, is self-evidently sound. There is no justification for modus ponens understood that way.

Although Frege indexes sentences and speaks of them as used in proof, the signs simply express and make formally perspicuous a proof structure of language- and mind-independent thought contents. Frege is very explicit that axioms are *Gedanken*, not symbolic expressions (*KS* 318–19/*CP* 334–5, "Über die Grundlagen der Geometrie iii" (1906), O 424; *NS* 221–2/*PW* 205–6, "Logik in der Mathematik" (1914)). The thought contents are epistemically fundamental. He intends to provide no justification for rules of inference understood as transitions among *thoughts*. Those transitions are self-evident.

So Frege never contemplates justifying axioms or rules of inference (in the fundamental sense of these terms, which apply to language-independent thought contents or rules), much less justifying them by semantical arguments that make reference to symbols. The idea of justifying the truth of axiomatic *Gedanken* by appeal to premises that refer to symbols, which are language- and mind-dependent, would have seemed absurd to him. So although the introduction of modus ponens in §14 is justified by a semantical soundness

against the formalists that his symbolism is not arbitrary. He is writing there about arithmetical truths (which certainly do, according to Frege, need justification—they are not basic truths), about the use of formulae, and about rules of inference as methods of moving from one formula (understood as having a sense) to another. Frege did not think of the axioms or language-independent rules as formulae at all. I can find no place where he speaks of justifying them. The language-independent axioms and rules are not in need of justification. Frege wants to justify his use of symbols, not the principles of inference underlying that use. Stanley makes substantially this point (p. 63). I do not, however, accept his claim that "Frege is treating his theory as an uninterpreted set of syntactic operations on strings of symbols". A semantics for meaningful language is still a semantics. Heck and Stanley are concerned to bring out the large role of semantical reasoning in *BL*. I think that Heck, in particular, is right in maintaining that all of the elucidations that lead to the introduction of the axioms are in effect part of semantical explanations of Frege's symbolism—even if they are equally explications of the relevant functions mentioned in the axioms.

argument, the justification is of an operation on his language, not of the underlying rule of inference among *Gedanken*. Although the introductions of the axioms are preceded, in some cases, by semantical explications of the terms in the expressions for the axioms, the axioms are not considered as true sentences, but as *Gedanken*. And the arguments for them are non-semantical arguments expressible within the logic. These arguments, as I have claimed, are articulations of the self-evidence of the axioms, which are not provable from more basic thoughts, and are not in need of proof. Both the rules of inference, as applying to thoughts, and the axioms (also thoughts) are taken to be self-evident. Only the language needs justification.

The issues over the sense in which Frege was doing semantics are complex and subtle. I think that Frege was clearly engaging in substantial and ineliminable semantical reasoning in various parts of *Basic Laws*. But I do not wish to discuss these issues further here. What is important for my purposes is that from an epistemic point of view, Frege was not taking the semantics of expressions to be more basic than his axioms, despite the fact that his work is clearly an early version of semantical reasoning. For the axioms whose formulations contain singular terms, he provides no argumentation at all, only an immediate appeal to an understanding of the axioms through the symbols that he has explicated. In the other cases, he does provide an argument with the axiom as conclusion. But these arguments are not semantical in the sense that they contain no steps that refer to symbols. And they do not argue from truths more basic than the conclusion. They do not justify, ground, or prove the conclusion in Frege's epistemically honorific sense of "proof". The arguments function, as the *Begriffsschrift* version of them clearly indicates, to provide an explication or articulation of the content of the axiom: they are discursive representations of an understanding of the axiom. The axioms, the basic true thoughts, are fundamental.

Let us look at this point more closely by returning to the argument for the first axiom in §18 of *Basic Laws*. Frege appeals to a line of the truth table for the conditional, a line that might be expressed (roughly) by this formula:

$$-(\Gamma \rightarrow \Delta) \rightarrow (\Gamma \,\&\, -\Delta).$$

Transforming this into a sentence relevant to the form of the axiom, we get:

$$-(a \rightarrow (b \rightarrow a)) \rightarrow (a \,\&\, (b \,\&\, -a)).$$

By non-contradiction, modus tollens, and double negation removal, we get:

$$(a \rightarrow (b \rightarrow a)).$$

This is one formalization of Frege's argument, but the formalization would not be the one Frege would give in his logic. For he explains "and" (§12) in terms of negation and the conditional. Thus

$$(\Gamma \text{ and } \Delta)$$

is explained as

$$-(\Gamma \to -\Delta).$$

Frege elsewhere clearly regards double negations as having *the same sense* as the result of dropping the double negations. It is not clear whether he regards the explication of "and" in terms of the negated conditional as giving the *sense* of sentences containing "and". But he may. He writes, "We see from these examples how the 'and' of ordinary language . . . [is] to be rendered" (§12).

Suppose that he *does* regard the "rendering" as giving the sense of the ordinary language sentence involving "and". Then the initial formula in the argument in §18 would have been conceived by Frege as having *the same sense* as

$$-(\Gamma \to \Delta) \to -(\Gamma \to \Delta).$$

The second step would have the sense of an analogously trivial truth, with its antecedent and consequent being identical. (Again we assume that double negation removal does not alter sense.) So the reasoning that Frege is articulating would not, on this view, be moving from one truth to another, but simply thinking through and expressing in different ways, via ordinary language, the character of the axiom which is the apparent conclusion of the reasoning. This seems to be his procedure in *Begriffsschrift* §14, though of course there he had not developed the sense–reference distinction.

This view of the sense of conjunctions would depend on a fairly (and I think implausibly) coarse-grained conception of the senses of logical expressions in ordinary language. But it is certainly not obvious that it was not Frege's view.

There is some circumstantial evidence for thinking that this might have been Frege's view of those arguments. In his writing on Axiom V outside of *Basic Laws*, he maintains that the two sides of the biconditional in the axiom *have the same sense* (*KS* 130–1; G&B 26–7, "Funktion und Begriff", O 10). Let us concentrate on instances of the axiom where the relevant function expressions are predicates. The two sides are

$$(x)(\mathrm{F}(x) \leftrightarrow \mathrm{G}(x))$$

and

$$\grave{\epsilon}\mathrm{F}(\epsilon) = \grave{\alpha}\mathrm{G}(\alpha).$$

Frege seems to have considered an argument that was supposed to bring out informally the supposed sameness of sense of these two sides, at least for the case in which "F" and "G" are predicates or concept expressions.

This argument, if it had been successful, would certainly have elicited the self-evidence of Axiom V, since the Axiom would have had the same sense as

$(x)(F(x) \leftrightarrow G(x)) \leftrightarrow (x)(F(x) \leftrightarrow G(x))$.

The argument would have justified Axiom V as a logical law, since the left side of Axiom V is certainly a proposition of pure logic, and the biconditional is a logical function.

Frege considered the case in which "F" and "G" are predicates or concept expressions. He maintained that "*x* is F" has the same sense as "*x* falls under the concept F". There is evidence that he thought that "*x* falls under the concept F" has the same sense as "*x* falls in the extension of the concept F". For "the concept F" *canonically* refers to an object—presumably an extension or course of values of the relevant predicate.[13] This sort of informal argument is something Frege appears to have been experimenting with in his writings on philosophy of language just before publication of *Basic Laws*. The unintuitive consequences of his view that "the concept F" denotes an object, not a concept, probably prevented him from articulating the argument in *Basic Laws*.

The argument would have carried out the same strategy that I have conjectured underlies the argument in §18 (for Axiom I). It is a manipulation of different ordinary expressions that have the same sense, but which purportedly brings out the self-evidence of a logical axiom.

It is not obvious, however, that Frege regarded his argument for Axiom I in §18 in the hyper-tautological way I have outlined. And it need not have been his view, for purposes of maintaining his position on the epistemic role of his arguments introducing the axioms. He indicates, as we have seen, that there are more basic truths in logical reality than are needed to axiomatize logic and arithmetic (*B* §13). So the reasoning in the semantical arguments in *Begriffsschrift* and *Basic Laws* might be seen as articulating the character of the self-evident axiom by appealing to an argument that one would have to be able to give in order to understand the axiom, even though the premises of the argument are no more basic than the axiom, and thus can provide no justification for it, or proof of it in Frege's epistemically freighted sense of "proof".

Understanding one basic truth may demand that one be able to understand its logical connections to others—even though understanding any one basic truth would suffice to recognize its truth and provide one with sufficient "evidence" to recognize its truth. (I will discuss this point further in §IV.) What the argument does is to bring out vividly the content of the axiom, whose understanding renders its truth evident.

It seems to me plausible that if one understands Axiom I, one realizes that it is true. One can by considering the condition under which it would be false realize that it cannot be false. Or one can simply recognize that it is true because it indicates that if *a* is true, it is true whether or not some arbitrary proposition is true. It seems to me that this understanding may necessarily

[13] I discuss this argument at greater length and provide citations in "Frege on Extensions of Concepts, From 1884 to 1903", *The Philosophical Review*, 93 (1984), 24–30; Ch. 7, 290–295.

require an ability to accept the argument that I outlined involving conjunction. Perhaps to understand the conditional one must understand conjunction (and vice versa). But the argument adds no justificatory force to a belief in the axiom. The belief is justified by the understanding involved in the belief. The argument provides no additional justification not already available from understanding the axiom. It does not proceed from steps more basic than the axiom. For to understand those steps, one must have an understanding of the material conditional sufficient to recognize the truth of the axiom independently of that particular argument.

These points simply instantiate Frege's view that axioms are basic truths, and basic truths do not need proof. Basic truths can be (justifiably) recognized as true by understanding their content.

I think that there are two interesting philosophical questions associated with the position that I have outlined. One is whether there are methods for determining a truth to be basic in the sense Frege relies upon. Frege gives no general account of such methods, and many would doubt that there is such a category of basic truths. But Frege does provide a range of philosophical arguments, particularly in *Foundations*, which are meant to persuade one of the basicness or non-basicness of various truths. Perhaps by systematizing and deepening the sorts of arguments Frege gives in *Foundations*, we could develop better ways for isolating such a category. We certainly have intuitions that some logical truths are more basic than others—and even some that seem as basic as possible. I am not persuaded that there is no fruitful subject matter here.

The other question is whether one can develop in more depth the relation between the point that understanding a thought requires inferential abilities (so that the understanding is articulable through inference), and the point that justification of belief in basic truths derives from understanding a single truth's content. The distinction is certainly tenable, but a fuller account of what goes into justification and what goes into understanding would be desirable.

Frege famously realized that understanding a thought requires understanding its inferential connections to other thoughts. So although a basic truth may carry its "evidence", its justification, in its content, an articulate understanding of that content might require connecting it to other contents, including truths that may be equally or less (but not more) basic. Frege's arguments "for" his axioms elicit understanding of the axioms by bringing out these connections.

Whether or not Frege understood his argument in §18 in the hyper-tautological way, the arguments that take his axioms as conclusions are compatible with his view that the axioms are basic truths that do not need or admit proof. The axioms are unprovable in that no genuine proof—no justification from more basic truths—is possible. Since they are self-evident, they do not need proof. They are self-evident because their justification is carried in their own contents. Understanding the content of an axiom suffices to warrant one in believing it. The point of the arguments is to articulate the content of the axioms and to elicit a firm understanding of them that resides in an understanding of their constituent senses.

III

Frege rarely supports his logical system or his logicism by invoking the traditional features of the axioms that we have discussed. Except for his insistence on rigour of formulation, much that is striking and original about his methodology bears little obvious relation to the Euclidean tradition. Although he alludes to self-evidence frequently, he almost never appeals to it in justifying his own logical theory or logical axioms. He never says or implies that convictions about self-evidence are infallible. I think that Frege believed that there is no infallible guarantee that one's commitments on logical or geometric truth are correct.

There is abundant evidence that Frege had some sympathy with modern caution about the reliability of appeals to what is obvious. He praised the refusal to be satisfied with even Euclidean standards of rigour, which refusal led to questioning the Parallel Postulate, and eventually to non-Euclidean geometries (*FA* §2). Although he thought that Euclidean geometries are true, he knew that he had eminent opponents. His belief that Euclidean geometry is true (indeed true of space) is surely based primarily on his sense that Euclidean geometry is more obvious than non-Euclidean geometries. But he bolsters this belief with argument: he argues that the geometries are incompatible, that Euclidean and non-Euclidean geometry cannot both be true, that a false system must be banned from the sciences, and that in view of its *longevity* one can hardly regard the Euclidean system as on a par with astrology. Frege uses these meta-considerations to support putting forward the axioms: "It is only if we do not dare to do this [treat Euclidean geometry on a par with astrology] that we can put Euclid's axioms forward as propositions that are neither false nor doubtful" (*NS* 184/*PW* 169, "Über die Euklidische Geometrie" (1899–1906?)). Although less than convincing, this is hardly the argument of a dogmatic rationalist.

Frege was not always conservative in his attitudes about traditional mathematical intuitions. He begins *Foundations of Arithmetic* with a litany of cases in which attempts to provide a foundation of proof had led to a sharper grasp of concepts, new mathematical theories, and deeper grounding of mathematical practice. He notes that epistemic standards in mathematics, especially in view of the advent of analysis, had been lax. He continues,

> Later developments, however, have shown more and more clearly that in mathematics a mere moral conviction, supported by a mass of successful applications, is not good

enough. Proof is now demanded of many things that formerly passed as self-evident ("*selbstverständlich*"). (*FA* §1)

He goes on to challenge the view that various arithmetic equations involving numbers in the thousands are self-evident ("*einleuchtend*") (*FA* §5).

Moreover, Frege backs Cantor's introduction of actual infinities, even though they were thought by many contemporary mathematicians to be deeply contrary to the limits of mathematical intuition (*FA* §85; *KS* 163 ff./ *CP* 178 ff., "Rezension von: Georg Cantor, Zur Lehre vom Transfiniten", O 269 ff.). He bases belief in the actual infinite not, of course, on direct mathematical intuitive powers, but on the role of the infinite in arithmetic and on his confidence that he could derive claims about it from arithmetical, and ultimately logical principles. But he clearly recognized that common mathematical beliefs about what is self-evident or intuitive or obvious could be flat out mistaken.

As I have noted, Frege's logic was not well-received by the dominant mathematicians of the day. They found not only his notation, but some of his principles misconceived. Frege responds not by insisting on the self-evidence of his principles, but by arguing that the only way to get a true logic is by providing a deeper analysis of judgments and inferential patterns than his Boolean opponents had provided (e.g. *NS* 37/*PW* 33, "Booles rechnende Logik and die Begriffsschrift" (1880–1881)). In *Basic Laws* we find Frege recommending to those who are sceptical of his logical system that they get to know it from the inside. He thinks that familiarity with the proofs themselves will engender more confidence in his basic principles (*GG* p. xii/*BL* 9; *FA* §90). In the Introduction to *Basic Laws*, Frege repeatedly appeals to advantages, to simplicity, and to the power of his axioms in producing proofs of widely recognized mathematical principles, as recommendations of his logical axioms.[14] He was aware that what people find intuitive or obvious is no safe guide to accepting or rejecting his own logical theory.

I think that we can assume that Frege thought that mathematical and logical intuition and judgment, even in outstanding mathematicians and logicians, is thoroughly fallible. Let me codify this point in two principles. He thought (a) that the fact that a mathematical or logical proposition is found obvious by competent professionals at a given time provides no infallible guarantee that it is true, much less a basic truth. He thought (b) that there is no guarantee that true mathematical or logical principles (including basic truths) will be found to be obvious by competent professionals at a given time.

The evidence for (b) is Frege's recognition of contemporary attitudes toward the axioms of Euclidean geometry and his awareness of scepticism

[14] This is the sort of argument usually associated with Ernst Zermelo ("A New Proof of the Possibility of a Well-Ordering", (1908) in J. van Heijenoort (ed.), *From Frege to Gödel* (Cambridge, Mass.: Harvard University Press, 1981)) in his defense of the axiom of choice. Frege's use of the argument form antedates Zermelo's.

about his own logical principles. The evidence for (a) is Frege's method of argument for accepting Euclidean geometry, his repeated criticism of "instinct" and "intuition" as ways of founding mathematics, his experience in struggling to find an acceptable logic against what was regarded by other (Boolean and Kantian) logicians as already constituting an acceptable logic, and perhaps his own uncertainty about Axiom V.[15]

This awareness of the fallibility of mathematicians' sense of what is obvious was part of the advanced spirit of the age. It did not constitute any sort of scepticism about mathematical knowledge, or even a concession that mathematical principles are less than "certain". Frege had a deep confidence in the ability of mathematical practice eventually to arrive at truth. And he maintained the traditional rationalist view that mathematics and logic are "certain" and epistemically more solid than empirical science. But these traditional views were tempered with a historical awareness of changes in these disciplines, and with an original thinker's awareness of how crooked the road to discovery could be. The nature of Frege's fallibilism will become clearer as we proceed.

Frege's method is non-Euclidean not only in his relative neglect of appeals to self-evidence when he is arguing for his logical theory, but also in his original way of developing that theory. As I have noted, in analyzing inferences Frege is concerned that appeals to self-evidence not be allowed to obscure the formal character of the inferences, which can be found only by rigorous logical analysis. This analysis is arrived at not primarily by consulting unaided intuition, but by surveying inferential patterns in actual scientific-mathematical reasoning. Frege ridicules the idea that one will find the appropriate logical concepts and logical structures ready-made by consulting intuition. He writes:

> All these concepts have been developed in science and have proved their fruitfulness. For this reason what we may discover in them has a far higher claim on our attention than anything that our everyday trains of thought might offer. For fruitfulness is the acid test of concepts, and scientific workshops the true field of study for logic. (*NS* 37/*PW* 33, "Booles rechnende Logik und die Begriffsschrift" (1880–1881))

As is well known, Frege's method was to reason to logical structure by observing patterns of judgments and patterns of inferences—and then postu-

[15] I have elsewhere discussed the bends and turns in Frege's changes of mind about Axiom V, and his attempts to persuade himself of its truth and its status as a basic law of logic ("Frege on Extensions of Concepts", 24–30; Ch. 7 above, pp. 290–295). It is clear, before as well as after his recognition of Russell's paradox, that Frege had and expected doubts about using the principle as an axiom, and seemingly even about its truth. Yet Frege did commit himself to Axiom V's being an axiom in the traditional sense, which would require that it be true, certain, unprovable, and self-evident—or not in need of proof. Any reflection on this situation at all would have enabled him to distinguish the objective property of self-evidence required of an axiom and his psychological state of finding the axiom less than completely obvious. He seemed to have hoped that the axiom would become more obvious with greater familiarity with the notion of an extension of a concept.

lating formal structures that would account for these patterns. While this method makes use of intuitions about deductive validity, it has at least as much kinship to theory construction as to intuitive mathematical reflection.

Frege's methods of analysis are closely associated with his famous contextualist methodological pronouncements. He holds that one can understand the "meaning" (later, sense and reference) of individual words only in the context of propositions. And he thought that one understands such semantical infrastructure only by understanding patterns of inferences—not by simple reflection. But understanding was traditionally supposed to be the basis for recognition of the truth of self-evident propositions. So if understanding requires such "discursive" procedures as logical analysis and theory construction—or at least the tacit abilities that such conscious construction codifies—it would seem that Frege's method constitutes a substantial qualification on traditional rationalist conceptions of reflection.

Frege's contextualism extends beyond methodology. He uses a contextualist argument for defending the existence of abstract objects in *Foundations of Arithmetic* (§§56–68.). He thinks that we are justified in believing in the existence of numbers as objects if we are justified in accepting mathematical propositions whose analysis shows number expressions to be singular terms.

On Frege's view, justification for accepting mathematical propositions seems to take three forms. Justification derives from what Frege calls "actual applications" (*FA* §§1–2). It derives from considerations of simplicity, duration, fruitfulness, and power in pure mathematical practice. It derives from understanding the self-evident foundations (axioms, definitions, inference rules) and from carrying out proofs.

Frege does not develop his notion of "actual applications" in detail. But the notion seems to attach to successful, applied mathematical practice. He appeals to the role of arithmetic or mathematical principles (like the associative law) in inductively supported applications within natural science or ordinary counting as one sort of application (*FA* §§2, 26). Here there seems to be an inductive confirmation of arithmetic through its success in application to non-mathematical domains.

The invocation of actual applications should be seen in the context of a broader conception of "pragmatic" justification within mathematical practice. Frege emphasizes that pure mathematical practice *works*. It produces a community of agreement through finding some systems "better", "simpler", "more enduring". It is this practice that Frege appealed to in defending Euclidean geometry, in the argument we discussed above.

Pragmatic considerations also enter into Frege's conception of the justification of definitions.[16] He thought that definitions are confirmed by their fruitfulness, by their ability to further mathematical practice (*FA* §88; *KS* 245/

[16] I hope that it is clear that by calling epistemic considerations "pragmatic" I am in no way implying that Frege thought them any less able to put us on to truth about a reality that is independent of our practice.

CP 255, "Über die Zahlen des Herrn H. Schubert", O 7). This view is nowhere more strikingly expressed than at the point of the key definition in *Foundations of Arithmetic*. Immediately after defining "the Number which belongs to the concept F", he writes, "That this definition is correct will perhaps be little evident ('*wenig einleuchten*') at first" (*FA* §69). He then goes on to argue that certain ordinary language objections to the definitions can be laid aside, because there is "basic agreement" between definiens and definiendum on our "basic assertions" about numbers and because discrepancies in ordinary usage do not raise serious rational objections to the definitions—only objections of habit and usage.

Frege entitles the next section "Completion and Proving-Good of our Definition". He writes, "Definitions prove good through their fruitfulness" (*FA* §70). Proving fruitful consists in aiding in a chain of proofs. Thus although as we have seen (*KS* 263/*CP* 274, "Über die Grundlagen der Geometrie i" (1903), O 320; *KS* 289–90/*CP* 302, "Über die Grundlagen der Geometrie" (1906), O 303) Frege held that definitions are self-evident (*selbst-verständlich*), and even, sometimes, that they preserve the sense of the definiendum (so that the senses of definitions are of the form "$a = a$"), he says here (*FA* §70) that a correct definition may be "little evident at first". They are self-evident in themselves, but not evident "at first" to us. Their intrinsic self-evidence might become more obvious to us over time, through their role in proofs; and we may receive some confirmation of their worth (self-evidence) through this role. Once the definitions are fully mastered and thoroughly used, presumably this external form of confirmation would be overdetermined and unneeded.

Instead of reflection, Frege appeals to mathematical practice—to observing the role of the definition in facilitating proof—as a way of confirming the worth, and seemingly the correctness, of the definition. Frege seems to have thought that one arrives at good definitions partly through the process of logical analysis, dependent on theory construction, that we have just been discussing.[17]

Frege thought that success in proving principles that are independently regarded as valid provides some justification for the principles (as well as definitions) used in carrying out the proof. Comparisons of simplicity and success in carrying out such proofs are "tests" of a system. He writes,

> The whole of the second part [the Proofs of the Basic Laws of Number] is really a test of my logical convictions. It is prima facie improbable that such a structure could be erected on a base that was uncertain or defective. Anyone who holds other convictions has only to try to erect a similar structure upon them, and I think he

[17] Frege's account of non-stipulative definition, and perhaps his conception of it, changes over the course of his career. But as far as I can see, these variations do not affect the primary points I am making. For a discussion of these matters which I think overrates the changes, but with which I am in basic agreement, see Dummett, *Frege: Philosophy of Mathematics*, (chs. 3 and 12).

will perceive that it does not work, or at least does not work so well. As a refutation in this I can only recognize someone's actually demonstrating either that a better, more durable edifice can be erected upon other fundamental convictions, or else that my principles lead to manifestly false conclusions. (*GG* p. xxvi/*BL* 25)

Here durability and working well, or working better, are tests of the basic principles and methods of his theory. It appears that Frege regarded these considerations as indicative of a kind of justificational support for his theory.

Frege saw these sorts of justification through "applications" and through "pragmatic" or other "methodological" considerations as insufficiently satisfying. They provide only "inductive", or only "prima facie", "probable" support. He shows little interest in justification through applications at all. Such justification might suffice to silence scepticism—in that one would find that the "boulder" is in fact "immovable" (*FA* §2). But it does not show what is holding the boulder so securely in place. Justification through consideration of pure mathematical practice seems to produce, for Frege, a secondary test, or a basis for prima facie perhaps inductive, justification. But Frege thought that the deeper justification lies in the structure of proof, which eventually leads back to logical axioms. He thought that the full "certainty", the rational unassailability, of mathematics would not be understood unless this proof structure was laid bare. Frege holds that whether he is right about his views about the logicist nature of this proof structure is a matter that can be determined only through the carrying out and the checking of the proofs (*FA* §90).

But assessing his logicism would require more than this. Knowing the order of reasons could not derive simply from checking proofs or soundness. It would also require finding that the proofs produce insight into justificational priority. The proofs in Frege's logical theory must match the proof structure that constitutes the justificational ordering among mathematical propositions—from first principles, including basic truths, to derivative principles. Such insight must derive from discursive reasoning both within and about the system. Frege thinks that recognizing that such a match has been attained is dependent on becoming familiar with his system, which is not just a matter of immediate reflection or insight. This view emerges repeatedly in the Introduction to *Basic Laws*.

Frege writes, "Because there are no gaps in the chains of inference, every 'axiom' . . . upon which a proof is based is brought to light; and in this way we gain a basis upon which to judge the epistemological nature of the law that is proved" (*GG* p. vii/*BL* 3). He continues:

I have drawn together everything that can facilitate a judgment as to whether the chains of inference are cohesive and the buttresses solid. If anyone should find anything defective, he must be able to state precisely where, according to him the error lies: in the Basic Laws, in the Definitions, in the Rules, or in the application of the Rules at a definite point. If we find everything in order, then we have accurate knowledge of the grounds upon which each individual theorem is based. (*GG* p. vii/*BL* 3)

Here again we encounter the view that one might find proposed axioms and inference rules in the foundations to be defective. The language clearly suggests that defect might in principle lie not only in the ordering or in proposed axioms not being logical, but even in proposed basic principles not being true or sound. "Finding everything in order" (which includes not only freedom from defect but being in the right justificational order) seems to require thorough familiarity with the system, and considerable discursive reasoning within it.

In recommending his new system Frege says that the introduction of courses of values of functions provides "far greater flexibility" and cannot be dispensed with (*GG* pp. ix–x/*BL* 6). He recommends the introduction of truth-values (in terms of which extensions of concepts are explained) by saying

How much simpler and sharper everything becomes by the introduction of truth-values, only detailed acquaintance with this book can show. These advantages alone put a great weight in the balance in favor of my own conception, which may seem strange at first sight. (*GG* p. x/*BL* 7)

Frege does not appeal to immediate insight or obviousness. He appeals to "advantages" which can be appreciated only by a detailed mastery of his theory, only through discursive reasoning.

Frege is aware of the unobviousness of his proposals:

I have moved farther away from the accepted conceptions, and have thereby stamped my views with an impress of paradox. . . . I myself can estimate to some extent the resistance with which my innovations will be met, because I had first to overcome something similar in myself in order to make them. (*GG* p. xi/*BL* 7)

Frege's view that acceptance of his proposals depends on detailed mastery of the system extends even to the acceptance of his *basic principles*, the basic axioms, and rules of inference:

After one has reached the end in this way, he may reread the Exposition of the *Begriffsschrift* as a connected whole. . . . In this way, I believe, the suspicion that may at first be aroused by my innovations will gradually be dispelled. The reader will recognize that my *basic principles* [my emphasis] at no point lead to consequences that he is not himself forced to acknowledge as correct. (*GG* p. xii/*BL* 9)

The basic principles gain something from our seeing what obviously correct consequences they have and from recognizing "advantages" of simplicity, sharpness, and the like. Here again, Frege is defending not merely the logicality but the truth of his basic proposed principles, through pragmatic modes of reasoning.

What do the basic principles gain from our seeing their consequences and our realizing their various "advantages"? If they are indeed axioms, they can be recognized as true "independently of other truths". The sort of justification that derives from understanding them and recognizing their truth through this understanding needs no further justificatory help from reflecting on their consequences or the advantages of the system in which they are embedded. The recognition of advantages seems to provide a prima facie, probabilistic justification that applies to the whole system, but derivatively to elements in it. Such recognition may provide indirect grounds for believing that the axioms are indeed basic and indeed true. But the supposed self-evidence of the axioms is ideally the primary source of their justification. They do not need or admit of proof. They gain a secondary, broadly inductive justification. And we gain greater sharpness of our understanding of them. I shall develop these points in §IV.

Let me summarize what I have been saying about Frege's epistemology. Frege thought that we are fallible in our convictions even on matters of self-evidence. He thought that we have no direct intuitive access to numbers as objects. Instead, we are justified in believing in them only through their role in making-true mathematical propositions that we are justified in believing. Our justification for believing in these propositions is partly pragmatic—we find their place in mathematical practice secure through long usage, through advantages of simplicity, plausibility, and fruitfulness, and through applications to non-mathematical domains. A deeper justification for believing in these propositions lies in finding their place in a logicist proof structure, by understanding the grounds within this structure that support them (if they are non-basic) or by understanding the self-evidently true basic principles.

Understanding this structure requires some of the discursive reasoning that plays a role in secondary, "pragmatic" justification. Understanding requires not only the logical analysis involving theory construction that I noted above, but also the production of proofs and the recognition that these proofs capture an antecedent order of justificational priority. Understanding the axioms requires, in some cases, reasoning from them in producing proofs, and even reasoning to them from reflection on their content. Frege knew, from painful experience, that other competent mathematicians would not immediately recognize his success (if he were indeed successful). He says that such recognition would depend on working through the proofs and acquiring increasing familiarity with the "advantages" that the conceptions that he had introduced offer. Thus whatever role self-evidence plays in his epistemology seems to be qualified by pragmatic considerations that result from reasoning within and about his system of proofs over time. I want to go into these qualifications in more detail in the next section.

IV

How does Frege's rationalist appeal to self-evidence accord with the fallibilist, pragmatic elements in his position? How could he appeal to pragmatic and philosophical considerations in persuading others of the analyticity of arithmetic and of the soundness of his logical system, when he held that justification ultimately comes down to self-evidence? Are the two tendencies simply ill-matched, ill-thought-out philosophical strands in the thinking of a mathematician?

This last question suggests serious underestimation of Frege's philosophical depth. I think that the integration of the two strands is one of his finest philosophical achievements. To make progress on our questions, we must scrutinize what Frege meant by "self-evident".[18]

Frege probably did not regard himself as using a well-honed technical term. Although no one has remarked on it, as far as I know, the term "self-evident" that appears in the standard English translations does not translate a single German counterpart. Sometimes Frege uses "*einleuchtend*" (and grammatical variants); sometimes he uses "*selbst-verständlich*"; and occasionally he uses "*evident*" and "*unmittelbar klar*". There are differences of meaning among these terms in colloquial German, but I have not found consistent differences in Frege's usage.[19]

I shall investigate the meaning of these terms for Frege by considering how he used them. Frege regarded both axioms of geometry (*FA*, combining §§13 and 90 ("*einleuchtend*"); cf. *NS* 183/*PW* 168, "Über die Euklidische Geometrie" (1899–1906?)) and logical axioms (*FA* §90 ("*einleuchtend*"); *GG* ii. §60/G&B 164 ("*selbst-verständlich*"); *GG* 253/*BL* 127 ("*einleuchtend*") as self-evident. At *KS* 393/*CP* 405, "Gedankengefüge", O 50, Frege writes, "the truth of a logical law is immediately evident ('*einleuchtet*') from itself, from the sense of its expression". By "logical law" he means a basic truth (or an axiom) of logic. (Cf. *KS* 262/*CP* 273, "Über die Grundlagen der Geometrie i" (1903), O 319.) Frege also regards rules of inference (*FA* §90 "*einleuchtend*") and thoughts expressed by propositions formed from correct definitions (*KS* 263/*CP* 274, "Über die Grundlagen der Geometrie i" (1903), O 320;, *KS* 289–90/*CP* 302, "Über die Grundlagen der Geometrie" (1906), O 303 ("*selbst-verständlich*") as self-evident. Axioms and thoughts expressed by propositions formed from correct definitions could be used as primitive steps in proofs; and rules of inference could be used as primitive modes of transition in proofs.

[18] I have benefited in this section from correspondence in 1991 with Robin Jeshion, who had thought independently about Frege's Euclideanism, and who emphasized the importance of not neglecting subjective elements in Frege's remarks about self-evidence.

[19] I do believe that "*selbst-verständlich*" is never used with any mentalistic overtones, whereas the others occasionally are. I will develop the matter of mentalistic overtones below.

It is not clear to me whether Frege regarded *any* truths of arithmetic as self-evident.[20] Very likely he did not. In *Foundations of Arithmetic* § 5 he criticizes those who take propositions involving addition of larger numbers as self-evident and demands proof of such propositions. Although he thinks that propositions involving addition of smaller numbers, like $1 + 1 = 2$, are *provable* and do not depend for their justification on Kantian intuition, he neither denies nor affirms that they are immediately self-evident ("*unmittelbar einleuchtend*" or "*unmittelbar klar*"). In so far as Frege was thinking of self-evidence in terms of recognizability as true independently of recognition of other truths (in the deepest proof-theoretic order of reasons), it is not surprising that he would not count arithmetic truths, even simple ones, as self-evident. (Cf. *KS* 393/*CP* 405, "Gedankengefüge", O 50.) It would seem that he might reserve the notion for basic truths, basic rules of inference, and definitions. At any rate, I know of no place where Frege counts arithmetical truths self-evident. There is reason to believe that Frege may have reserved the notion of self-evidence for truths that are "not in need of proof".

Frege does, however, regard arithmetical truths as *certain*, that is, beyond reasonable doubt given understanding of the proposition (*B* § 13; *FA* §2). And in § 14 of *Foundations* Frege says that denying any of the fundamental propositions of arithmetic leads to complete confusion: "Even to think at all seems no longer possible". This point serves the view that arithmetic has a wider domain than geometry—the domain of everything thinkable. But he regards this argument for the analyticity of arithmetic as non-demonstrative (*FA* §90). So he regarded his claim that denying fundamental arithmetical principles would throw thought into complete confusion as fallible.

Our discussion in §I shows that self-evidence is not the same as obviousness, or immediate psychological certainty. In Frege's primary usage, self-evidence appears to be compatible with lack of obviousness to individuals. Frege thought that something could be an axiom and yet be found by professional mathematicians or logicians to be unobvious. He speaks of propositions that "formerly passed as self-evident" ("*selbst-verständlich*") (*FA* §1). Here he implies that something might seem to be self-evident but not be so.

Frege himself was uncertain about at least one of the thoughts that he proposed as an axiom (hence as self-evident). The Introduction to *Basic Laws* concludes with a statement that fatefully leaves open the possibility that one of the basic principles, one of the proposed axioms, is defective; he even identified the faulty "axiom" as the only likely source of difficulty. Since he thought it was an axiom, he must have, at least sometimes, thought that *it* was certain, but because of insufficient analysis or incomplete understanding, *he*

[20] One could imagine such exceptions as $1 = 1$, which are themselves obvious logical truths. I mean arithmetical truths whose truth depends essentially on arithmetical notions.

was not.[21] Moreover, he found that other basic principles of his logic were not universally accepted by opposing logicians. But he maintained the views, which he several times expresses in the pre-paradox period, that the basic principles that he proposed are genuine axioms and that axioms are self-evident.

The distinction between self-evidence and psychological certainty, or felt obviousness, is a corollary of the distinction between justification and discovery that Frege draws at the opening of *Begriffsschrift* (quoted in §I). Self-evidence on one primary construal is meant to be bound up with the account of the justification of logical and mathematical truths, not with psychological means of discovering their truth.

The foundations of logic and mathematics were supposed to be "unprovable"—not justifiable by derivation from other thought contents. But Frege thought that one was nevertheless justified in holding them to be true. Their justification rests not only on their being unprovable, but on their not being "in need of proof". That is, they are rationally acceptable in themselves. This is part of the literal etymological meaning of "self-evident": they carry their "evidence", their rational support, in themselves and are dependent on none from other propositions.

Frege states this point quite directly: "...it is part of the concept of an axiom that it can be recognized as true independently of other truths" (*NS* 183/*PW* 168, "Über die Euklidische Geometrie" (1899–1906?)). This remark

[21] Richard Heck ("Frege and Semantics") holds that Frege was in doubt only about the status of Axiom V as a logical truth, not about its truth. Although this is not obviously false, I see no clear evidence for this view, and some evidence to the contrary. Frege probably did worry about its logical status independently of worrying about its truth. But I believe that he was uneasy about its truth as well. It was clearly less obvious than the other axioms, because the notion of *course of values* was relatively new to mathematics, despite its connection to the notion of extension and the graph of a function. The dispute between Frege and the iterative set theorists suggests that there was fundamental doubt on both sides about the viability of the notions (respectively) of set and of extension. Cantor was not committed to, and probably would not have accepted, Axiom V. Frege, who had read Cantor, would have been aware of this. (Cf. Burge, "Frege on Extensions of Concepts" (Ch. 7 above).) Moreover, Frege's concerns about Axiom V in the introduction of *BL* explicitly focus on its truth—as well as its status: He invites the reader to find "error" or something "defective" in (among other places) *BL* (*GG* p. vii/*BL* 3), and he explicitly indicates that a dispute can arise only with respect to Axiom V. He associates this invitation with dispelling suspicion of his principles by having the reader see that they lead to no mistakes:

> In this way, I believe, the suspicion that may at first be aroused by my innovations will gradually be dispelled. The reader will recognize that my *basic principles* [my emphasis] at no point lead to consequences that he is not himself forced to acknowledge as correct. (*GG* p. xii/*BL* 9)

And he ends the introduction by again raising the possibility of error, falsehood—as well as the possibility that different axioms would produce a more durable structure. He declares (with fatefully exaggerated bravado) that no one will be able to find such error (*GG* p. xxvi/*BL* 25). As Heck points out, Frege took himself to have provided a meta-theoretic "justification" of Axiom V in §31. But Frege could not have thought that a semantical proof had any more certainty than the axiom that it explicated or justified, since if Axiom V was an axiom in Frege's sense, it was self-evident and did not need proof. Thus any psychological uncertainty about the axiom (uncertainty expressed *after* the giving of the proof) would have to transfer to uncertainty about the proof as well. Given that the proof used new methods (and in fact turned out to be defective), it is not surprising that Frege would have retained doubts even after giving the proof.

is virtually echoed in terms of self-evidence: "...the truth of a logical law [a basic truth of logic] is immediately evident from itself, from the sense of its expression" (*KS* 393/*CP* 405, "Gedankengefüge", O 50). Frege does not mean that it is psychologically possible to recognize it as true independently of other thoughts. For the concepts of an axiom and of a basic law are explicitly (in many places) intended to be independent of psychological considerations. But the appeal to recognizability obviously involves some implicit reference to mind, or a recognizing capacity. I think that this reference is to an ideal rational mind with full understanding.

The implicit reference to mind is also contained in Frege's basic epistemic concerns. Frege is interested in the sources or springs of knowledge. He announces his interest in knowledge in the opening passage from *Begriffsschrift* that I quoted at the outset of this paper. The basic laws are laws on which all *knowledge* rests. In the middle of his career (*KS* 221/*CP* 235, "Über die Begriffsschrift des Herrn Peano und meine Eigene" (1896)) and at the end (*NS* 286–94/*PW* 267–74, "Erkenntnisquellen der Mathematik und der mathematischen Naturwissenschaften" (1924/1925)), he remains interested in the springs of knowledge—empirical, logical, geometrical. Knowledge involves a contribution by mind—belief or judgment. To understand knowledge, one must conceive of some relation between the purely abstract proof structure, the propositional system of grounds or reasons, and mind.

Similarly, in framing his account of his basic epistemic categories—apriority and analyticity—Frege has in mind some conception of justification for judgment (holding-true), not simply justification or grounding for an abstract proposition or thought content:

> Now these distinctions between a priori and a posteriori, synthetic and analytic, concern, as I see it, not the content of the judgment but the justification for making the judgment.... When a proposition is called a posteriori or analytic in my sense, this is not a judgment about the conditions, psychological, physiological and physical, which have made it possible to form the content of the proposition in our consciousness; nor is it a judgment about the way in which some other person has come, perhaps erroneously, to believe it true; rather, it is a judgment about the ultimate ground upon which rests the justification (*Berechtigung*) for holding it to be true. (*FA* §3)

The ultimate justificatory ground (the ground or the justification) is independent of minds. The content of any mathematical justification is independent of minds. The problem of the foundation of arithmetic is basically a problem in mathematics. But that ground or content is justification for *belief*, or holding-true, or recognition of truth.

The relevant notion of mind here though is abstract and ideal. There is no reference to individual minds or to the psychology of recognition, belief, or judgment. When Frege writes, "...it is part of the concept of an axiom that it can be recognized as true independently of other truths" (*NS* 183/*PW* 168, "Über die Euklidische Geometrie" (1899–1906?)), he means that the truth can be rationally and correctly recognized as true by a rational mind independently of resting the rationality of this recognition on derivation of the truth from other recognized truths.[22]

The basic truths are laws at the foundation of a justificational structure. The other truths receive their justification by being logically derivative from the basic ones. And the basic ones carry their justification intrinsically, in that their truth can be justifiably recognized from the nature of those truths, in justificational independence of consideration of other truths. On this conception self-evidence is an intrinsic property of the basic truths, rules, and thoughts expressed by definitions. It is intrinsic in that it is independent of relations to actual individuals. It does involve implicit relation to an ideal mind. But ideal minds are abstractions, themselves understood in terms of rational capacities, which are characterized in terms of the abstract justificational structures. These structures are for Frege independent of individual minds, or indeed any individuals.

Although Frege thought that axioms (basic truths) *can* be recognized as true and basic by *actual human minds*, it is certain that he did not think that it is part of the concept of an axiom (or basic truth) that they can be. In so far as it is a necessary truth that axioms are recognizable as true by human minds, it would be a truth that derived from necessary conditions on any possible mind (*qua* mind), not from conditions placed on axioms or basic truths by the notions of actual mind or human mind. In most of the formulations containing such notions as "unprovability", "certainty", "recognizability", Frege intends implicit reference to some sort of ideally rational mind. His notion of an ideally rational mind is, as I have emphasized, constitutively dependent on the abstract structures that define the norms for rationality.

I have outlined a rational reconstruction of what might be called Frege's objective conception of self-evidence, one that accords with much in the rationalist tradition. To give a rough summary: A truth is self-evident in this sense if (i) an ideally rational mind would be rational in believing it; (ii) this rationality in believing it need not depend for its rationality on inferring it from other truths—or reasoning about its relation to other truths; it derives merely from understanding it; and (iii) belief in it is unavoidable for an ideally rational mind that fully and deeply understands it. I think that this is the main thrust of what Frege intended in requiring that axioms, rules of inference, and definitions be self-evident.

But this cannot be a complete account of Frege's conceptions of self-evidence. Even if justification is supposed to derive somehow from an

[22] Cf. also *GG* ii. §147/G & B 181. These references to an ideal rational mind are not psychologistic, since there is no reference to actual minds, and since the notion of a rational mind is understood in terms of a capacity to recognize the truth of and logical relations between elements in the Platonic logical structure that Frege is discussing. Cf. Burge, "Frege on Knowing the Third Realm", *Mind*, 101 (1992), 633–49; Ch. 8 above.

abstract structure that is independent of minds and waiting to be discovered, it is clear that Frege is not indifferent to the fact that *individuals* are justified in accepting truths in this structure. He does not ignore actual knowledge or actual justifications for individuals. He is quite aware that *we* have mathematical and logical knowledge, and that *we* make judgments. In order to discuss how his abstract normative principles bear on actual mathematical practice and knowledge, Frege had to relate the abstract structure of thoughts (or propositions) to minds. Through most of his career, Frege has little to say about such a relation. But some points can be extracted for discussion here.

Frege's terms that translate "self-evident" usually make no explicit reference to actual minds. But there are some uses that explicitly or implicitly presuppose such reference. The word "*einleuchtend*" translates normally as obvious, clear, or evident. The verb from which the adjective derives takes the dative in many standard constructions. Frege typically uses the term without the dative (often intending implicit relation to some ideal rational mind), but at least once he includes the dative (*FA* §90): "Often . . . the correctness of such a transition is immediately self-evident *to us*" [my emphasis]. As we have seen (in *FA* §69) he notes that his primary definition will be "little evident at first"—clearly intending that it will be little evident to ordinary individual reasoners, at first.

Moreover, (in *FA* §5) Frege argues from the fact that certain thoughts are not immediately clear (*klar*) or immediately evident (*einleuchtend*) to the conclusion that they are not axioms but are provable. He explicitly regards this not as a demonstrative argument, but as one that adds probability (*FA* §90). If he meant "self-evident" purely in the sense of "self-justifying", he would be arguing that since they are in need of justification by reference to other propositions, they are not axioms and thus are in need of proof (and since they are justified, they are provable). But this *would* be a demonstrative argument, and its premise would beg the question against his opponents.[23] Clearly, Frege's premise must be taken to be a remark about the fact that it is not immediately (non-inferentially) obvious to our minds that the relevant thought contents are true.

Similarly, when in 1903 Frege learns that Axiom V is false, he remarks that he had always recognized that it is not as self-evident (*einleuchtend*) as the others (*GG* ii. 253/*BL* 127). The comparative form suggests obviousness to actual minds, which admits of degrees. Frege also says that he always

[23] Frege could have been arguing from the assumption that it is not obvious to an ideally rational and understanding mind, to the conclusion that it is not an axiom. And he might have regarded the assumption as "merely probable". This argument seems enthymemic. It requires the additional premise that we are (probably) ideally rational and that we have ideal understanding of the arithmetic propositions. But there are two reasons for thinking that this is not Frege's argument. In the first place, Frege clearly thought that before him no one had full understanding of the arithmetic propositions, since no one had uncovered their logical natures. In the second, Frege appears to be appealing to immediate intellectual experience of our not finding large addition problems immediately obvious—not to some circuitous argument in terms of ideal rationality.

recognized that Axiom V lacks the self-evidence that must be required of logical laws. Here too he must be saying that it had not achieved the obviousness to actual minds (his mind) that should underwrite the postulation of thought contents as logical laws. It would have been obviously untrue for him to have said that he always recognized that Axiom V did not meet the objective requirement of ideal self-evidence that is a requirement on logical laws. For then he would have been involved, before the discovery of the paradox, in an obvious and easily discovered inconsistency.

If he had always recognized Axiom V not to be self-evident in the more objective sense that axioms are required to be self-evident, he would never have proposed it as an axiom, as a first principle at the foundation of the structure of logic and mathematics. For doing so would be obviously inconsistent with his own principles. Indeed, in *Basic Laws* he says "it must be demanded that every assertion that is not completely self-evident should have a real proof" (*GG* ii. §60; G & B 164). Axiom V is certainly asserted without "real proof"—a deductive argument from more basic truths. Frege was surely not involved in such a simple-minded inconsistency before the discovery of the paradox. His use of "self-evident" ("*einleuchtend*") in the appendix seems to have a somewhat different meaning from the meaning it has in his official pronouncements on the self-evidence of axioms. Perhaps Frege is sliding between more and less objective uses of "self-evident". But the issue is very subtle and requires further elaboration.

Before discovering the paradox, Frege recognized that Axiom V was less obvious to him than the other axioms. But he must have thought—at least as the public position that he committed himself to—that it *was in fact* self-evident in a sense required of axioms, a sense that is compatible with not being fully obvious or self-evident to him or other individuals. He must have thought that it is compelling for a fully rational mind that fully understands the principle. In finding it less than fully obvious, he must have regarded himself as having a less than ideal insight into the principle. After discovery of the paradox he realized that the lack of obviousness had turned out to be a sign of trouble with the proposition, not merely insufficiency in his insight.

How are we to take Frege's remark in the appendix to *Basic Laws* that self-evidence, in a sense that includes a relatively subjective component, is required of axioms? In that remark he seems to say that axioms must be required to be self-evident in a sense that would include (in addition to other things) a large degree of obviousness to individuals, whereas Axiom V always lacked the requisite degree of obviousness to individuals.

Now it is possible that one should regard Frege's remark as autobiographically important, but insignificant for casting light on his reflective views. For it was written in a hurry and at a time of anguish, immediately on being informed by Russell of the contradiction in his system.

But the same tension between objective requirements of self-evidence and subjective experiences of obviousness for axioms emerged in the Euclidean tradition. For example, some regarded the axiom of parallels as insufficiently (subjectively) obvious to be an axiom. Others took this lesser obviousness as a historical fact or a human failing which did not bear on the objective foundational status of the axiom—its self-evidence in the sense of being recognizable as true without needing or admitting proof.

There is good reason for inquiry to require that proposed axioms be obvious in a subjective sense. For if a proposition is not immediately obvious, then one will want either further justification of it—some sort of proof or further argument—or additional clarification of one's understanding of it. And this state might lead one reasonably not to take it as a starting point in justification.

On the other hand, it seems that a requirement that axioms be obvious to individuals, though having a subjective component, must retain some element of objectivity. It cannot be a final requirement on axioms or basic truths, only a deeply significant prima facie requirement. It must allow for individuals' learning histories. Every sophisticated rationalist allowed that truths that are self-evident could be foggily grasped or apprehended. So the fact that an individual does not find some proposition obvious cannot in general show that the proposition is not self-evident, or that it cannot be taken as the foundation for a justificational system. So subjective obviousness cannot be an absolute requirement on an objectively correct system. At any time an individual can be unsatisfied with a principle as an axiom because of its lack of obviousness. Whether this feature is a final objection to the principle's being an axiom depends on whether the lack lies in the principle or in the individual. Still, inevitably, one must make use of what is obvious in trying to determine what is (ultimately) objectively self-evident. One must place some reliance on one's own contingent and perhaps limited rationality and understanding.

How then are we to understand the relation between the intrinsic self-justifying character of the axioms and the clarity or obviousness of their truth to *us*?

Answering this question depends on understanding the enormous degree to which Frege thought we tend to understand concepts, or thought components, incompletely. He thought that in arithmetic and logic, we think with thought components whose correct explication or real inferential structure we do not fully understand. His argument for his logic presupposes, and often makes explicit, that most logical categories are obscured to us by the non-logical functions of language. His argument for his definitions of arithmetical terms, in *The Foundations of Arithmetic*, is prefaced with and saturated by an emphasis on the idea that only through centuries of labour (culminating in his own work) had mankind come to a true understanding of the concept of number and of various other arithmetical notions (*FA* p. vii).[24] Frege

[24] For extensive discussion of this issue in Frege, see Burge, "Frege on Sense and Linguistic Meaning", in D. Bell and N. Cooper (eds.), *The Analytic Tradition* (Oxford: Blackwell, 1990), Ch. 6 above; *idem*, "Frege on Extensions of Concepts", Ch. 7 above.

recognizes that the correctness of his critical definition "will hardly be evident at first" (*FA* §69—another psychologically tinged usage). And he goes on to try to make it "evident" by improving our understanding of the critical notions involved in it. Definitions that are self-evident may not *seem* to be self-evident, but they may help us become more clearly aware of the content (sense) of a word that we have already been using.

In *Basic Laws* Frege indicates that the introduction of truth-values (which are basic for his explanation of terms in Axiom V) makes everything "sharper" in a way that only detailed acquaintance with the book can show (*GG* p. x/*BL* 7—quoted in section II). Sharpness is Frege's term for complete grasp or understanding. (Cf. footnote 24.) Understanding the basic principles is enhanced by seeing their role in deriving theorems of the system (*GG* p. xii/*BL* 9).

What is it to be self-evident but not seem so? Descartes' appeals to self-evidence were not meant to receive immediate approbation from anyone who would listen. Self-evident propositions are self-evident to anyone who adequately *understands* them. Descartes and other traditional rationalists did not assume that understanding would be immediate or common to all mankind, or even common to all socially accepted experts. Understanding was something more than mere mastery of the words or concepts to a communal or conventional standard. It involved mastering a deeper rational and explanatory order. Acknowledgment of the truth of self-evident propositions would be immediate and non-inferential, only given full understanding in the relevant deeper sense.

Frege is relying on this tradition. This reliance explains his belief that our acceptance of thoughts as basic principles is fallible. It explains his view that familiarity with the details of his system will enhance acceptance of his basic principles. Full understanding is necessary for the self-evidence of basic principles to be psychologically obvious (evident, clear).

What is original about his position is not his view that a thought might be self-evident but not seem self-evident—self-evident but not obvious to an individual. It is not his idea that subjective obviousness or subjective unobviousness might submit to reversal through deeper conceptual development and understanding. What is original is his integration of these traditional views with his deep conception of what goes into adequate understanding. This conception rests on his method of finding logical structure through studying patterns of inference. Coming to an understanding of logical structure is necessary to full understanding of a thought. And understanding logical structure derives from seeing what structures are most fruitful in accounting for the patterns of inference that we reflectively engage in.

Thus Frege's "pragmatic" claim, that one will see the correctness and lose the sense of strangeness of his first principles by reasoning within his system and seeing the various advantages yielded by his proposals, is compatible with his appeal to self-evidence. It serves not to justify the first principles (except in a secondary, inductive way which will be overshadowed, given full understanding) but to engender full understanding of them. One might recognize the truth of the axioms independently of other truths only in so far as one fully understands the axioms. But understanding them depends not only on understanding Frege's elucidatory remarks about the interpretation of his symbols, but also on understanding their logical structure—their power to entail other truths, and their reason-giving priority. This latter understanding is *not independent* of reasoning that connects them to other truths. All full understanding involves discursive elements, even if recognition of the truth of axioms is, given sufficient understanding, "immediate".

Frege's view that incomplete understanding might impede obviousness of self-evident principles also explains why he thought that philosophical argumentation was worthwhile in helping people see that his axiomatic theory matches a proof structure that constitutes the order of justification in logic and mathematics. Like mathematical practice, philosophical argumentation deepens understanding. By associating arithmetical principles with general features of logic—chiefly universality of subject matter and relevance to understanding norms of thought and judgment—Frege hoped to improve our understanding of the kinship between logic and arithmetic, and sharpen our sense of logical, conceptual, epistemic priority. By distinguishing between sense and tone, sense and reference, function and object, he hoped to clarify our understanding of first principles.

As I have noted, I think that Frege's pragmatism and contextualism play another, secondary, role in his epistemology. They not only play a role in accounting for understanding. They provide a secondary, fallible, non-demonstrative justification. Frege seems to think that reflection on mathematical practice—which is hardly separable from much of what he counted philosophy—provides supplementary strong grounds to accept his definitions and his exposition of logical priority, even though a foundational proof structure, when fully understood, provides deeper grounds. This justification, which dominates *The Foundations of Arithmetic*, seems to be independent of the carrying out of proofs that display proof-theoretic justifications. It is hard to see coming to understand justificational priority of a truth, as opposed to mere mathematical acceptability of a proof, apart from such reflection on mathematical practice. Although Frege regarded justification implicit in the shape, stability, and fruitfulness of mathematical practice as secondary (and presumably he did not think of it as *strengthening* proof-theoretic justifications), it seems to have been for him a first defence against scepticism, and a sufficient one. I am inclined to think that his confidence on this score is the main reason why he never developed his conception of the relation between human understanding and self-evidence.

10 *Frege on Apriority (2000)*

Frege's logicism incorporated both a set of purported proofs in mathematical logic and an investigation into the epistemology of arithmetic. The epistemological investigation was for him the motivating one. He saw his project as revealing 'the springs of knowledge' and the nature of arithmetical justification. Frege maintained a sophisticated version of the Euclidean position that knowledge of the axioms and theorems of logic, geometry, and arithmetic rests on the *self-evidence* of the axioms, definitions, and rules of inference.[1] The account combines the traditional rationalist view that beliefs that seem obvious are fallible and understanding is hard to come by, with his original insistence that understanding depends not primarily on immediate insight but on a web of inferential capacities.

Central to Frege's rationalism is his view that knowledge of logic and mathematics is fundamentally apriori. In fact, near the end of *The Foundations of Arithmetic* he states that the purpose of the book is to make it probable that 'the laws of arithmetic are analytic judgements and consequently apriori.'[2] In this essay I want to discuss Frege's conception of apriority, with particular reference to its roots in the conceptions of apriority advanced by Leibniz and Kant.

Frege advertised his notion of apriority as a 'clarification' of Kant's notion. It is well known that Frege did not read Kant with serious historical intent. But even allowing for this fact, his advertisement seems to me inter-

I gave a shorter version of this paper at a conference on Frege in Bonn, Germany, in October 1998. I am indebted to John Carriero, Wolfgang Künne, Christopher Peacocke, Rainer Stuhlmann-Laisz, and Christian Wenzel for comments that led to improvements. In the text I use 'apriori' and 'aposteriori' as single English words. In quotations from Frege, Kant, and Leibniz, I use the Latin phrases 'a priori' and 'a posteriori'. I use 'apriority' instead of the barbaric 'aprioricity'. The latter is the misbegotten result of drawing a mistaken parallel to 'analyticity'. It would be appropriate to the form 'aprioric' which of course has no use. 'Aprioricity' bears no natural relation to 'apriori'. Quine pointed this out to me several years back.

[1] Burge, "Frege on Knowing the Foundation", *Mind*, 107 (1998), 305–47; Ch. 9 above.

[2] *FA* §87. (Translations are mine. I have consulted Austin's free but often elegant renderings. I will henceforth cite this book by section under the abbreviation "*FA*" in the text.) Frege's view of analyticity has been more often discussed than his view of apriority. Essentially he takes a proposition to be analytic if it is an axiom of logic or derivable from axioms of logic together with definitions. He rejects conceptions of analyticity that would tie it to emptiness of substantive content.

estingly misleading. I believe that his notion is in important respects very different from Kant's and more indebted to Leibniz.

I

Frege's only extensive explication of his conception of apriority occurs early in *The Foundations of Arithmetic*. He begins by emphasizing that his conception concerns the ultimate canonical justification associated with a judgement, not the content of truths:

These distinctions between a priori and a posteriori, synthetic and analytic, concern, as I see it,* not the content of the judgement but the justification for the judgement-pronouncement [*Urteilsfällung*]. Where there is no such justification, the possibility of drawing the distinctions vanishes. An a priori error is thus just as much a non-entity [*Unding*] as a blue concept. When a proposition is called a posteriori or analytic in my sense, this is not a judgement about the conditions, psychological, physiological, and physical, which have made it possible to form the content of the proposition in our consciousness; nor is it a judgement about the way in which another has come, perhaps erroneously, to believe it true; rather, it is a judgement about the deepest ground upon which rests the justification for holding it to be true.
*(Frege's footnote): By this I do not, of course, want to assign a new sense but only meet [*treffen*] what earlier writers, particularly Kant, have meant. (*FA* §3)

Frege writes here of apriori judgements. But afterwards he writes of apriori propositions and then apriori truths, and eventually (*FA* §87, cf. n. 1) apriori laws. These differences are, I think, not deeply significant. Judgements in Frege's sense are idealized abstractions, commitments of logic or other sciences, not the acts of individuals. Individuals can instantiate these judgements through their acts of judgement, but the abstract judgements themselves seem to be independent of individual mental acts. Truths and judgements are, of course, different for Frege. But the difference in Frege's logic concerns only their role in the logical structure. Some truths (true antecedents in conditionals) are not judged. They are not marked by the assertion sign. But everything that is judged is true.[3]

[3] This doctrine is of a piece with Frege's view that (in logic) inferences can be drawn only from truths. Here he means not that individuals cannot infer things from falsehoods, but rather that the idealized inferences treated in logic proceed only from true axioms. Inferences for Frege are steps in proof that constitute ideal, correct justifications that exhibit the natural justificatory order of truths. Michael Dummett seems to me to get backwards Frege's motivations for the view that proofs have to start from true premisses and that one should not derive a theorem by starting with a supposition. Dummett claims that Frege believed that a complete justification is possible *because* of his rejection of inference from reductio or from other suppositions. Cf. Michael Dummett, *Frege: Philosophy of Mathematics* (Cambridge, Mass.: Harvard University Press, 1991), 25–6. It seems clear that Frege rejected such inference because he thought of proofs as deductive arguments that reveal natural justificatory order. I do not see that this view is incompatible with allowing 'proofs' *in the modern sense* that proceed without axioms, by natural deduction. Frege's conception of proof is very different from the modern one. It is concerned with an ideal, natural order of justification. Leibniz also thought

Only truths or veridical judgements can be apriori for Frege. He writes that an apriori error is as impossible as a blue concept. Frege justifies his claim that only truths can be apriori by claiming that apriority concerns the nature of the justification for a judgement. Of course, some ordinary judgements can be justified without being true. But Frege seems to be focused on justifications—deductive proofs from self-evident propositions—that cannot lead judgement into error. Here Frege signals his concern with canonical, ideal, rational justifications, for which the truth-guaranteeing principles and proofs of mathematics and logic provide the paradigm.

In predicating apriority of truths and judgements, understood as canonical commitments of logic and mathematics, Frege is following Leibniz.[4] Leibniz gave the first modern explication of apriority. He maintained that a truth is apriori if it is knowable independent of experience.[5] Since Leibniz explicitly indicates that one might depend psychologically on sense experience in order to come to know any truth, he means that a truth is apriori if the justificational force involved in the knowledge's justification is independent of experience.

Like Leibniz, Frege conceives of apriority as applying primarily to abstract intentional structures. Leibniz applied the notion not only to truths but to proofs, conceived as abstract sequences of truths.[6] Frege assumes that all justifications are proofs, indeed deductive proofs. The apriority of a justification (a series of truths constituting a deductive argument) resides in the character of the premises and rules of inference—in the 'deepest ground [justification, reason, *Grund*] on which a judgement rests'. Like Leibniz, Frege thinks that there is a natural order of justification, which consists in a natural justificatory order among truths.

Frege's definitional explication of apriority continues directly from the passage quoted above:

> Thus the question is removed from the sphere of psychology and referred, if the truth concerned is a mathematical one, to the sphere of mathematics. It comes down to finding the proof and following it back to the primitive truths. If on this path one comes only upon general logical laws and on definitions, one has an analytic truth,

of reductios as second-class proofs because they do not reveal the fundamental order of justification. Cf. Leibniz, *New Essays on Human Understanding* (New York: Cambridge University Press, 1705, 1765, 1989), III. iii. 15; also R. M. Adams, *Leibniz: Determinist Theist, Idealist* (Oxford: Oxford University Press, 1994), 109–10. I believe that Dummett may be right in holding that Frege's actual mathematical practice may have been hampered by too strict a focus on the justificatory ideal.

[4] Frege makes explicit his dependence on Leibniz on these matters in Frege (*FA* §17): 'we are concerned here not with the mode of discovery but with the ground of proof; or as Leibniz says, "the question here is not about the history of our discoveries, which differs in different men, but about the connection and natural order of truths, which is always the same." ' Frege draws his quotation of Leibniz from Leibniz, *New Essays*, IV. vii. 9.

[5] Cf. Leibniz, *New Essays*, IV, ix, 1, 434; Leibniz, *Philosophical Essays*, ed. R. Ariew and D. Garber (Indianapolis: Hackett Publishing Co., 1989): "Primary Truths", 31; "On Freedom", 97.

[6] Cf. Leibniz *Philosophical Essays*, "First Truths", and "Monadology", §45. There are also occasional attributions of apriority to knowledge or acquaintance. But Leibniz is fairly constant in attributing apriority primarily to truths and proofs.

bearing in mind that one must take account also of all propositions upon which the admissibility of a definition depends. If, however, it is impossible to carry out the proof without making use of truths which are not of a general logical nature, but refer to the sphere of a special science, then the proposition is a synthetic one. For a truth to be a posteriori, it must be that its proof will not work out [*auskommen*] without reference to facts, i.e. to unprovable truths which are not general [*ohne Allgemeinheit*], and which contain assertions about determinate objects [*bestimmte Gegendstände*]. If, on the contrary, it is possible to derive the proof purely from general laws, which themselves neither need nor admit of proof, then the truth is a priori. (*FA* §3)[7]

Several points about this famous passage need to be made at the outset. Frege distinguishes apriori from analytic truths. Analytic truths derive from general *logical* truths together perhaps with definitions. Apriori truths derive from general laws, regardless of whether they are logical, which neither need nor admit of proof. It is clear that Frege regards all analytic truths as apriori. The relevant logical truths are laws. It is also clear that Frege regards some apriori truths as not analytic. Truths of geometry are synthetic apriori, thus not analytic.

Although Frege uses the language of giving a sufficient condition in explaining 'analytic', 'synthetic', and 'apriori', he uses the language of giving a necessary condition in the case of 'aposteriori'. But if Frege elsewhere (e.g. *FA* §63) uses 'if' when he clearly means 'if and only if'. And I believe on a variety of grounds that Frege sees himself as giving necessary and sufficient conditions, probably even definitions, for all four notions. I will mention some of these grounds.

In the passage from section 3 that precedes this passage, and which I quoted above, Frege states that judgements concerning all four categories are judgements about 'the deepest ground upon which rests the justification for holding [the relevant proposition] to be true'. This is a characterization of the nature and content of the judgement, not merely a characterization of a necessary or a sufficient condition. The subsequent individual characterizations of the four categories characterize the nature of the deepest grounds for a predication of each category. Thus they seem to be characterizing the nature of the judgements predicating the categories ('analytic', 'apriori', and so on), not merely one or another condition on their truth.

Frege's characterization of analytic judgements is clearly intended to provide necessary as well as sufficient conditions. His characterization is traditional and may be seen as a gloss on Leibniz's view of analyticity. The parallel in language between the characterization of analytic judgements and

[7] Austin's translation mistakenly speaks of giving or constructing a proof, which might suggest that the definition concerns what is possible for a human being to do. In fact, Frege's language is abstract and impersonal. His account concerns the nature of the mathematical structures, not human capacities. Austin translates 'bestimmte Gegenstände' as 'particular objects'. In this, I think that he is capturing Frege's intent, but I prefer the more literal translation.

the characterization of apriori judgements counts towards regarding both as necessary and sufficient. As we shall see, the characterization of apriori judgements is one that Leibniz held to be necessary and sufficient, at least for finite minds. This was known to Frege. If he had meant to be less committal than Leibniz on the matter, it would have been natural for him to have said so.

In the fifth sentence of section 13 of *Foundations*, Frege tries to reconcile his account of intuition as singular with the view that the laws of geometry are general. Frege believed that the laws of geometry are general, that they are about a special domain of (spatial) objects, and that they do not need or admit of proof. He is assuming that arithmetic and geometry are both apriori, but that they differ in that geometry rests on intuition and is synthetic. I think that Frege is attempting to deal with the worry that since geometrical laws rest on intuition, they must rest on particular facts. This would not be a worry unless Frege regarded it as a necessary as well as sufficient condition on the apriority of geometry that its proofs rest on general laws that neither need nor admit of proof.

Frege's solution is roughly that intuition in geometry does not make reference to particular objects, and the geometric proofs begin with self-evident general principles, though they rest in some way, not well articulated (and which we shall discuss later), on the 'not really particular' intuition. Frege is clearly trying to explicate intuition's role in geometry in a way that leaves that role compatible with taking geometrical proofs to be grounded purely in general geometrical laws. Presumably intuition is meant to be part and parcel of the self-evidence of the general laws.

I do not think that Frege quite solves the problem, particularly inasmuch as he intends to agree with Kant. As I shall later argue, Frege is trying to marry a Leibnizian conception of apriority with Kant's account of synthetic apriority in geometry while siding with Leibniz about the analyticity of arithmetic. The marriage is not a complete success.

There is, of course, a parallel worry that runs as an undercurrent through Frege's mature work. The worry is that the attempt to derive arithmetic from general logical laws (which is required for it to be analytic apriori) is incompatible with the particularity of the numbers. How does one derive particularity from generality in arithmetic? The issue is signalled at the end of *Foundations*, § 13. This worry centers ultimately on Law V, which is the bridge in Frege's mature theory from generality to particularity. Again, this issue is naturally seen in the light of the demand that in being analytic apriori, arithmetic must derive ultimately from logical laws that are purely general.

For these and other reasons, I shall assume that Frege's characterization of apriority in *Foundations*, § 3 is intended as a necessary and sufficient condition. In fact, I think he views it as a definition.

The notion of a fact about a determinate object in Frege's explication of aposteriori truth in the passage from § 3 is reminiscent of Leibniz's identification of aposteriori truths with truths of 'fact', as contrasted with truths of reason.[8] Frege and Leibniz agree in not seeing truths of reason as any less 'factual' than truths of fact. The point is not that they are not factual, but that they are not 'merely' factual, not merely contingent happenstance. They are principles that are fundamental or necessary to the very nature of things. The point that apriori truths are *general* is basic to the Leibniz–Frege conception of apriority. I will return to this point.

Frege departed from Leibniz in thinking that apriori truths include both truths of reason and synthetic apriori truths that involve a combination of reason and geometrical spatial intuition. In this, of course, Frege follows Kant. I believe, however, that Frege's departure from Leibniz on this point is not as fully Kantian as it might first appear. I shall return to this point as well.

Mill had claimed that all justification ultimately rests on induction.[9] Turning Mill virtually on his head, Frege holds that empirical inductive justification is a species of deductive proof, which contains singular statements together with some general principle of induction as premises (*FA* § 3). He does not make clear what he considers the form of the deduction to be. And he does not indicate in his definition of aposteriori truths how he thinks singular judgements about 'facts' are justified. Presumably he thinks the justification depends in some way on sense experience. It seems likely that he regarded sense-perceptual observations of facts as primitively justified aposteriori. For our purposes, it is enough that Frege thought that justifications relevant to apriori truths are either deductive proofs or self-evident truths. Such justifications have to start with premises that are self-evident and general.

Frege assumed that all apriori truths, other than basic ones, are provable within a comprehensive deductive system. Gödel's incompleteness theorems undermine this assumption. But insofar as one conceives of proof informally as an epistemic ordering among truths, one can perhaps see Frege's vision of an epistemic ordering as worth developing, with appropriate adjustments, despite this problem.[10]

Frege writes that the axioms 'neither need nor admit of proof'. This phrase is indicative of Frege's view of proof as a canonical justificational ordering of truths, or ideal judgements, that is independent of individual minds or theories. Any truth can be 'proved' within some logical theory, in the usual modern sense of the word 'prove'. But Frege conceived of proof in terms of natural or canonical justification. He saw some truths as fundamental 'unprovable' truths, axioms, or canonical starting points in a system of ideal canonical

[8] Cf. e.g. Leibniz, *New Essays*, IV. vii. 9. 412; IV. ix. 1. 434.

[9] J. S. Mill, *System of Logic* (1843) (New York: Harper & Bros., 1893), II. VI. 1.

[10] Michael Dummett, *Frege: Philosophy of Mathematics*, 29–30, in effect makes this point. Dummett errs, however, in thinking that Frege is concerned with what is knowable by *us* (cf. ibid. 24, 26, 28–9). There is no such parameter in Frege's account. The natural order of justification among truths is conceived as a matter that is independent of whether *we* can follow it.

justification. Such primitive truths do not need proof in that they are self-evident or self-justifying. And they cannot be justified through derivation from other truths, because no other truths are justificationally more basic. Thus they do not admit of proof in his sense. The formula of basic truths and axioms neither needing nor admitting of proof can be found *verbatim* in Leibniz, from whom Frege surely got it.[11]

In introducing his conception of apriority, Frege follows the traditional rationalist practice of indicating the *compatibility* of apriority with various sorts of dependence on experience. In particular, Frege notes that a truth can be apriori even though being able to think it, and learning that it is true, might each depend on having sense experience of facts.[12] Whether a truth is apriori depends on the nature of its canonical justification. Thus one could need to see symbols or diagrams in order to learn a logical or mathematical truth. One could need sense experience—perhaps in interlocution or simply in observing various stable objects in the world—in order to be able to think with certain logical or mathematical concepts. Perhaps, for example, to count or to use a quantifier, one needs to be able to track physical objects. But these facts about learning or psychological development do not show that the propositions that one thinks, once one has undergone the relevant development, are not apriori. Whether they are apriori depends on the nature of their justification. Frege thinks that such justification in logic and mathematics is independent of how the concepts are acquired, and independent of how individuals come to recognize the truths as true.

In his discussion of Mill's empiricism, Frege reiterates the point:

> If one calls a proposition empirical because we have to have made observations in order to become conscious of its content, one does not use the word 'empirical' in the sense in which it is opposed to 'a priori'. One is then making a psychological statement, which concerns only the content of the proposition; whether the proposition is true is not touched. (*FA* § 8)[13]

[11] Leibniz, *New Essays*, e.g. IV. ix. 2, 434. The formula also occurs in Lotze. Perhaps Frege got the phrase from Leibniz through Lotze. Cf. R. H. Lotze, *Logik* (Leipzig, 1880); *idem*, *Logic*, trans. B. Bosanquet (Oxford, 1888; repr. New York, 1980), § 200. Frege seems, however, to have read Leibniz's *New Essays*. I discuss this notion of proof and Frege's view of axioms in some detail in Burge. "Frege on Knowing the Foundation", *Mind*, 107 (1998), 305–47; Ch. 9 above.

[12] Cf. this section of the passage quoted above: 'When a proposition is called a posteriori . . . in my sense, this is not a judgement about the conditions, psychological, physiological, and physical, which have made it possible to form the content of the proposition in our consciousness; nor is it a judgement about the way in which another man has come, perhaps erroneously, to believe it true; rather, it is a judgement about the deepest ground upon which rests the justification for holding it to be true.'

[13] Frege fixes here on truth, not justification. I think that he is assuming that one learns something about the nature of apriori truths by understanding the proof structure in which they are embedded; and this proof structure constitutes their canonical justification. Cf. Frege, *FA* § 105.

Substantially the same distinction between the nature of a truth (and ultimately its justification) and the ways we come to understand the relevant proposition or to realize its truth is made by Leibniz, *New Essays*, Preface, 48–9; IV. vii. 9; and by Kant, *Critique of Pure Reason* (1781, 1787), A1/B1.

The key element in the rationalist approach is this distinction between questions about the psychology of acquisition or learning and normative questions regarding the nature of the justification of the propositions or capacities thus learned.

I say 'propositions or capacities'. Frege follows Leibniz in predicating apriority of propositions, or more particularly, truths, or sequences of truths—*not* capacities, or mental states, or justifications associated with types of propositional attitudes. Apriority ultimately concerns justification. But Leibniz and Frege share the view that apriority is a feature of an ideal or canonical way of justifying a proposition. For them, a proposition is either apriori or aposteriori, but not both, depending on the nature of the ideal or canonical justification associated with it.[14]

In this, Leibniz and Frege diverge from one distinctive aspect of Kant's thinking about apriority. Like Leibniz and Frege, Kant predicates apriority in a variety of ways—to intuitions, concepts, truths, cognition, constructions, principles, judgements. But whereas Leibniz and Frege predicate apriority primarily of truths (or more fundamentally, proofs of truths), Kant predicates apriority primarily of cognition and the employment of representations. For him apriori cognition is cognition that is justificationally independent of sense experience, and of 'all impressions of the senses'.[15] Apriori cognition is for Kant cognition whose justificational resources derive purely from the function of cognitive capacities in contributing to cognition. Apriori employment of concepts (or other representations) is employment that carries a warrant that is independent of sensory experiences. Aposteriori cognition is cognition which is justificationally derivative, in part, from sense experiences.

Both conceptions are ultimately epistemic. Frege very clearly states that his classification concerns 'the ultimate ground on which the justification for taking [a truth] to be true depends' (*FA* § 3). Both sharply distinguish epistemic questions from questions of actual human psychology. Both take apriority to hinge on the source or method of warrant.

One might think that the main difference lies in the fact that Kant acknowledges more types of warrant as sources of apriority. Leibniz and Frege allow self-evidence and proof. Kant allows, in addition, constructions that rest on pure intuition and reflection on the nature of cognitive faculties.

I think, however, that this difference is associated with a fundamentally different orientation towards apriority. Frege and Leibniz explicate the nature of apriority in terms of a deduction from general basic self-evident truths. All

[14] Some modern philosophers who take apriority to be predicated primarily of propositions call a proposition apriori if it *can* be justified apriori. Apriori justification is then explained in some non-circular way. Cf. S. Kripke, *Naming and Necessity* (Cambridge, Mass.: Harvard University Press, 1972), 34. This formulation avoids commitment to that way's being canonical or ideal. But it also leaves out a serious commitment of such rationalists as Leibniz and Frege. For them, apriori justification is the best and most fundamental sort of justification. When something can be known or justified apriori, that is the canonical way.

[15] Kant, *Critique of Pure Reason*, B2–3.

that matters to apriority is encoded in the eternal, agent-independent truths themselves. For deductive proof turns entirely on such contents. An individual's being apriori justified consists just in thinking through the deductive sequence with understanding.

For Kant, the apriority of mathematics depends on possible constructions involving a faculty, pure intuition, that does not directly contribute components of truths, the conceptual components of propositions or thoughts. According to Kant, the proofs in arithmetic and geometry are not purely sequences of propositions. The justifications, both in believing axioms and in drawing inferences from them, must lean on imaginative constructions in pure intuition, which cannot be reduced to a sequence of truths. The intuitive faculty contributes singular images in apriori imagination. Not only are these not part of an eternal order of conceptual contents. The proofs themselves essentially involve mental activity and make essential reference, through intuition, to particulars. For Kant these particulars are aspects of the mind. So the structure of a mathematical proof makes essential reference to possible mental particulars. It is not an eternal sequence of truths that are fundamentally independent of particulars.

Kant's conception of synthetic apriori cognition thus depends on an activity, a type of synthesis involved in the making of intuitive constructions in pure imagination. It is significant that, unlike Leibniz and Frege, he makes no appeal to self-evidence. That is, he does not claim that the evidence for believing the basic truths of geometry and arithmetic is encoded in the truths. In arithmetic he does not even think that axiomatic proof is the basis of arithmetical practice.[16]

This orientation helps explain Kant's tendency to predicate apriority of cognition rather than truths. It is also at the root of his concentration, in his investigation of apriori warrant, on the functions and operations of cognitive capacities, not on the nature of conceptual content and the relations among truths. The orientation makes the question of what it is to *have* a justification much more complex and interesting than it is on the Leibniz–Frege conception. And it ties that question more closely to what an apriori warrant is.

Kant's shift in his understanding of apriority from the content of truth and of proof-sequences of propositions to the character of cognitive procedures opens considerably more possibilities for understanding sources of apriority, and for seeing its nature in capacities and their functions, or even in specific acts or mental occurrences, rather than purely in propositional forms. Kant's account does not depend on empirical psychology, but it does center on a transcendental psychology of the cognitive capacities of any rational agent.[17]

[16] Kant, *Critique of Pure Reason*, A164/B205.

[17] The relation between the two approaches is complex and needs further exploration. But it is worth remarking that Kant's approach has this advantage of flexibility. For Leibniz and Frege, a truth is either apriori or aposteriori. It is apriori if its canonical or ideal mode of justification is apriori. Its canonical mode of justification is apriori if it is situated in a natural proof structure either as a primitive

A second way in which Frege diverges from Kant is that his explanation of apriority in *The Foundations of Arithmetic*, §3, makes no mention of sense experience. Instead he characterizes it in terms of the generality of the premises of its proof.[18] Both Leibniz and Kant characterize apriority directly in terms of justificational independence of experience. Unlike Leibniz, Kant consistently takes experience to be *sense* experience. Since any modern notion of apriority seems necessarily tied somehow with justificational independence of experience, Frege's omission is, strictly speaking, a mischaracterization of the notion of apriority.[19]

From one point of view, this omission is not of great importance. Frege evidently took his notion of apriority to be equivalent with justificational independence of sense experience. His discussion of Millian empiricism follows his definition of apriority by a few pages. In those sections he

truth—which does 'not need or admit of proof'—or as a deductive consequence from primitive truths and rules of inference. On Kant's conception, a truth can be known or justifiably believed either apriori or aposteriori, depending on what form of justificational procedure supports it for the individual. For on this conception, apriority is predicated not primarily of truths but of modes of justification, or even states of cognition. Kant did not make use of this flexibility. Its possibility is, however, implicit in his conception.

Michael Dummett, *Frege: Philosophy of Mathematics*, 27, writes: 'it is natural to take Frege as meaning that an a priori proposition may be known a posteriori: otherwise the status of the proposition would be determined by any *correct* justification that could be given for it.' He goes on to discuss whether there are any propositions that can be known only apriori. I have no quarrel with Dummett's substantive discussion. But his historical reasoning is off the mark. Frege's characterization takes apriority to apply to truths or idealized judgements. There is no relativization to particular ways of knowing those truths. A truth or judgement-type is either apriori or not. A truth or judgement is apriori if its best or canonical justification proceeds as a deductive proof from general principles that neither need nor admit of proof. Dummett fails to notice that there is no clear meaning within Frege's terminology for a question whether a truth can be known both apriori and aposteriori. That question can be better investigated by shifting to a Kantian conception of apriority. Dummett slides between the two conceptions. Frege could certainly have understood and accepted the Kantian conception; but he did not use it or propose it.

Dummett's reasoning to his interpretation is unsound. Suppose for the sake of argument that we reject the view that an apriori proposition can be known aposteriori. (I myself would resist such a rejection.) We might allow that there are empirical justifications for something weaker than knowledge for all propositions. For example, we might strictly maintain the Leibniz–Frege conception and insist that apriori truths can be known only apriori. Then it simply does not follow that the status of the proposition would be determined by any correct justification that could be given for it. The status would still be determined by the best justification that could be given for it. Oddly, Dummett clearly sees that this is Frege's conception elsewhere—ibid. 23.

[18] As Dummett notes, Frege's definition of 'apriori' is cast in such a way that the premisses of apriori proofs are counted neither apriori nor aposteriori (ibid. 24). I think that Dummett is correct in thinking this an oversight of no great significance. It would be easy and appropriate to count the primitive truths and rules of inference apriori.

[19] There are differences between Leibniz's and Kant's accounts on this point that are relevant, but which I intend to discuss elsewhere. Leibniz often characterizes apriority in terms of justificational independence of experience. Leibniz sometimes allows intellectual apprehension of intellectual events to count as 'experience'. Kant firmly characterizes apriority in terms of justificational independence of *sense* experience. Kant's specification has important consequences, and makes his view in this respect the more modern one. It was taken up by Mill, the positivists, and most other twentieth-century empiricists. For purposes of epistemological discussion, 'experience' has come to mean *sense experience*.

repeatedly writes of 'observed facts', apparently picking up on the notion of fact that appears in his definition of aposteriority (*FA* §§7–9). He seems to assume that mere 'facts'—unprovable truths that are not general—can enter into justifications only through observation.[20] So a proof's depending on particular facts would make it rest on sense experience. Moreover, his criticism of Mill explicitly takes 'empirical' to be opposed to 'apriori' (*FA* §8).

In *The Foundations of Arithmetic*, section 11, Frege infers from a proof's not depending on examples to its independence of 'evidence of the senses'. The inference suggests that he thought that a proof from general truths necessarily is justificationally independent of sense experience. At the beginning of *The Basic Laws of Arithmetic* he states the purpose of *The Foundations of Arithmetic* as having been to make it plausible that arithmetic is a branch of logic and 'need not borrow any ground of proof whatever from either experience or intuition'.[21] Here also Frege assumes that a proof's proceeding from general logical principles entails its justificational independence from experience or intuition. Frege commonly accepts the Kantian association of intuition (in humans) with sensibility, so here again it is plausible that he meant by 'experience' 'sense experience'.

In very late work, forty years after the statement of his definition, Frege divides sources of knowledge into three categories: sense perception, the logical source of knowledge, and the geometrical source of knowledge. He infers in this passage from a source's not being that of sense perception that it is apriori.[22]

So Frege took his definition of apriority in terms of derivation from general truths to be equivalent to a more normal definition that would characterize apriority in terms of justificational independence from sense experience. Still, the non-standardness (incorrectness) of Frege's definition is interesting on at least two counts. First, its focus on generality rather than independence from sense experience reveals ways in which Frege is following out Leibnizian themes but in a distinctively Fregean form. Second, the definition is backed by a presupposition, shared with Leibniz, that there is a necessary equivalence between justifications, at least for finite minds, that start from general principles and justifications that are justificationally independent of sense experience. It is of some interest, I think, to raise questions about this presupposition.

[20] Precisely the same inference can be found in Leibniz, *New Essays*, Preface 49–50.

[21] *BL*, *GG*, § 0. Compare this characterization of the earlier book's purpose with the one quoted from Frege (*FA* §87; cf. note 1 above). It is possible that the latter characterization constitutes a correction of the mischaracterization of apriority in *FA* § 3.

[22] *PW* 267 ff., 276–7; *NS* 286 ff., "Erkenntnisquellen der Mathematik und der mathematischen Naturwissenschaften" (1924/1925); *NS* 296–7, "Zahlen und Arithmetik" (1924–1925).

II

Let us start with the first point of interest. Leibniz and Frege both see apriori truth as fundamentally general. Apriori truths are derivable from general, universally quantified, truths. Both, as we have seen, contrast apriori truths with mere truths of fact. Leibniz held that mere truths of fact are contingent, and that apriori truths are necessary. He took necessary truths to be either general or derivable from general logical principles together with definitional analyses and logical rules of inference. So for Leibniz the apriori–aposteriori distinction lines up with the necessary–contingent distinction, and both are closely associated with Leibniz's conception of a distinction between general truths and particular truths.[23]

It is tempting to regard Frege in the same light. As we have seen, Frege even defines apriority in terms of derivability from general truths and aposteriority in terms of derivability from particular truths. But there is little evidence that Frege associated apriority or generality with necessity. In fact, modal categories are strikingly absent from Frege's discussion.

We can gain a more refined understanding of Frege's differences from both Leibniz and Kant by contrasting his terminology with Kant's. Kant's conception of apriority, as we have seen, is explicitly defined in terms of a cognition's independence for its transcendental or epistemological genesis and its justification from sense experience. But he cites two other properties as marks (*Merkmale*) or sure indications (*sichere Kennzeichen*) of apriority. One is necessity. The other is strict generality (or universality) (*strenge Allgemeinheit*).[24]

[23] Leibniz, *New Essays*, Preface 49–50; IV. vii. 2–10, 408–13; IV. xi. 13, 445–6. The characterization of Leibniz's view that I use in this section, which brings it very close to Frege's, depends on laying aside Leibniz's views of God's cognition. The characterization seems to me true of Leibniz's view of finite, human cognition, but less obviously true of his view of divine cognition. Leibniz thought that God could have apriori knowledge of contingent truths through infinite analysis. By analysing the infinitely complex individual concepts of contingently existing individuals, God could know all truths about them. Cf. "Necessary and Contingent Truths", in *Leibniz: Logical Papers*, trans. and ed. G. H. R. Parkinson (Oxford: Clarendon Press, 1966), 97 ff. In discussing God's infinite analyses, Leibniz lays no explicit weight on the generality of the apriori truths. It is not clear that Leibniz thought that for God apriority is ultimately general. What Leibniz emphasizes is analysis of contents in such a way as to resolve them into identities. There is little discussion of the nature of the contents, where they come from, how they are determined. Whether these truths, which are knowable apriori through formal analysis by God, are ultimately singular is open to question. On the other hand, Leibniz thought that even individuals are reflections of a plan of God's. It therefore seems possible that contingent singular identities, on Leibniz's view, are ultimately instantiations of some general rational plan, which might have the status of a general law. Cf. *Discourse on Metaphysics*, § 6. Leibniz sees singular identities in logic and mathematics as resolvable into identities that instantiate general necessary truths. These are truths that are knowable apriori by finite minds. In this Leibniz and Frege are one. Regardless of whether he thought that a generalization lies at the bottom of infinite analysis of contingent truth, Leibniz is also kin to Frege in his emphasis on the idea that apriority lies in formal structure. I am indebted for these qualifications to John Carriero.

[24] Kant, *Critique of Pure Reason*, B3–4; cf. A2; A91–2/B124. The same point is made in Kant, *Critique of Judgment*, §7; *Akademische Ausgabe* (*AA*), V. 213. There Kant calls comparative gener-

There are two points to be noted about these remarks. One is that Kant provides these marks or indications not as elements in the definition of apriority, but as signs, which according to his theory are necessarily associated with apriority. In fact, in providing these signs, he takes them to be sufficient for apriority. He does not, in these famous passages, claim that they are necessary conditions.[25] The reason why on his view apriori judgements are associated with necessity and strict generality, is not that these associations follow from his definition or conception of apriority. The associations derive from further commitments in Kant's system.

Kant explains strict generality itself in terms of modality. Kant contrasts strict generality with comparative or assumed generality. Comparative generality holds only as far as we have observed.[26] A judgement thought in strict generality 'permits no possible exception'. Kant infers from this that such a thought is taken as holding absolutely apriori.

Neither Kant nor Leibniz gives any hint of defining apriority in terms of generality. Both appeal, however, to generality in their elucidations of apriority. Frege's use of generality (*Allgemeinheit*) in his definition is surely inherited from them. Like them he believed that apriority is deeply connected with some form of generality of application, or universal validity. But he interpreted and used his notion of generality differently. He departs from both Leibniz and Kant in *defining* apriority in terms of generality. He departs from

ality 'only general' (*nur generale*), and strict generality 'universal' (*universale*). Compare Leibniz, *New Essays*, IV. ix. 14, 446: 'The distinction you draw [between particular and general propositions] appears to amount to mine, between "propositions of fact" and "propositions of reason". Propositions of fact can also become general, in a way; but that is by induction or observation, so that what we have is only a multitude of similar facts . . . This is not perfect generality, since we cannot see its necessity. General propositions of reason are necessary.'

[25] I think that Kant believed that necessity is (necessarily) necessary as well as sufficient for the apriority of a judgement. He clearly believed that being, or being derivable from, a strictly general proposition is sufficient for the apriority of a judgement. Kant surely believed that all apriori judgements are true without any possible exceptions. Whether he believed that all apriori judgements have to be derivable from judgements that are in the form of universal generalizations is more doubtful. I shall discuss this matter below. Whether strict generality was only a sufficient condition (a mark) of apriority, not a necessary one—or whether it was both necessary and sufficient, but understood in such a way as not to entail the logical form of a generalization—is a complex question that I shall leave open. What is certain is that Kant's views on the relation between apriority and both necessity and strict generality depend not merely on his definition or conception of apriority, but on other elements in his system. I believe that rejecting Kant's positions on these relations is compatible with maintaining his conception of apriority.

[26] Cf. n. 24. Strictly speaking comparative generality and strict generality do not seem to be exhaustive categories. It would appear that there are propositions that are comparatively general but which are not true accidental generalizations (there is a counter-instance that simply has not been found); yet true accidental generalizations are not necessary truths. This is because it is possible for there to be true accidental generalizations which have no counter-instances yet observed. (I leave open whether there are also empirical laws which are general but which are not strictly general, in Kant's sense.) It is possible, of course, that Kant means the 'we' in 'what we have so far observed' in a loose and highly idealized sense. It is conceivable that he intended comparative generality to include all possible actual observations by 'us'. Given his idealism, he would take this as equivalent to the empirical truth of the generalization. This is a matter that could bear more investigation.

both in saying little about the relation between apriority and necessity. Indeed, his conception of generality differs from both in that he does not connect it to modal notions, seen as independent notions, at all.

Frege does comment on the relation between generality and necessity very briefly in *Begriffsschrift*. He associates generality with the logical form of the contents of judgements. He claims that apodictic judgements differentiate themselves from merely assertoric ones in that they suggest the existence of general judgements from which the proposition can be inferred. He then writes:

> When I designate a proposition as necessary, I thereby give a hint about the grounds of my judgement. But since the conceptual content of the judgement is not thereby touched, the form of the apodictic judgement has no significance for us.[27]

Frege seems to think that necessity is not represented in logical form, but is to be explained in terms of a pragmatic suggestion regarding the epistemic grounds for a judgement. Generality for Frege (in the sense relevant to this context) is simply universal quantification. What makes a truth apriori is that its ultimate grounds are universally quantified. So Frege seems to explicate necessity in terms of apriority. Apriority is the notion that Frege attaches in *Foundations of Arithmetic* to the condition he envisages here in the *Begriffsschrift* of a judgement's having its ground in general propositions. If anything, Frege explains necessity in terms of the (ordinary) generality of the grounds of the proposition. This contrasts with Kant's explaining (strict) generality in terms of necessity.

I think that Frege was trying to get the effect of the difference between accidental generalizations and empirical laws, on the one hand, and necessary generalizations, on the other, while avoiding explicit introduction of independent modal notions. His notion of generality is the simple one of universal quantification. Not just any general truth is apriori, however. Only general truths that are self-evident axioms, or first-truths, or which are derivable from self-evident axioms, or first-truths, are apriori. Apriori generalizations are generalizations whose *ultimate justification* does not rest on particular truths.

Frege does use the notion of law in his characterization of apriority: 'If . . . it is possible to derive the proof purely from general laws, which themselves neither need nor admit of proof, then the truth is apriori' (*FA* §3).[28] Empirical laws need and admit of 'proof', in that they need justification from statements of observation about particulars. It is common to hold that the notion of law contains or implies modal notions. That may well be. But I believe that Frege thought of laws in terms of basic principles in a system of scientific propositions—either an empirical science or a deductive

[27] *B* §4. The issue is discussed briefly by G. Gabriel, "Frege's 'Epistemology in Disguise' ", in M. Schirn (ed.), *Frege: Importance and Legacy* (Berlin: Walter de Gruyter, 1996).

[28] I believe that Frege's use of 'possible' in this remark is dispensable. It is possible to derive the proof in his sense if and only if there is a proof.

science—not (at least not officially) in terms of any modal or counterfactual element. Empirical laws are basic principles of idealized empirical scientific systems of true, grounded propositions. But they are not basic in the order of justification: singular observational statements (along with an apriori principle of induction) are supposed to be justificationally prior. Apriori laws differ in just this respect.

So the key idea in distinguishing empirical laws and accidental generalizations from apriori truths is taking apriority to be justificational derivation from general truths, which themselves are self-evident and do not need or admit of proof. Frege's notion of *generality* is fundamentally less modal than Kant's notion of strict generality or universality. It is simply that of universal quantification, where quantification is understood to be unqualifiedly general—to range over *everything*. Apriority is understood in terms of the priority of generality in justification.[29]

I have no doubt that Frege worked with an intuitive notion of logical validity. This enters his formulation of rules of inference. But the universal validity of *logical* laws is supposed to lie in their applicability to *everything*—which includes mathematical and geometrical objects and functions. The mathematical objects provide a sufficiently large and strict subject matter to enable true quantifications in logic and mathematics to have some of the force and effect of necessary truths that purport to quantify over possible objects or possible worlds. This force and effect seems to suffice for Frege's purposes. Frege seems to avoid invocation of an independent notion of modality and of merely possible objects, in epistemology, metaphysics, and logic.

Leibniz took all truths to be deducible in principle from truths of logic. On his view, it is a mere weakness of the finite human intellect that requires it to invoke empirical experience to arrive at ordinary truths about the physical world. Frege joined the rest of mankind in regarding Leibniz's view as overblown (*FA* §15). Of course, he agreed with Leibniz in holding that arithmetic is derivable from logic. Logic is naturally seen as a canon of general principles associated with valid inference. Here Frege sided with Leibniz against Kant in holding that one can derive truths about particular, determinate objects—the numbers—from purely general logical principles. Frege specifically states his opposition to Kant's view that without sensibility, no object would be given to us (*FA* §89). He argues that he can derive the existence of numbers from purely general logical laws. In this, of course, he failed. But the Leibnizian idea of obtaining truths about particular determinate objects from general, logical, apriori principles is fundamental to his logicist project.

[29] Frege has another concept of 'generality', of course, by which he distinguishes arithmetic and logic, which are completely general in their domain of applicability, from geometry, which applies only to space.

It seems to me likely that Frege's opposition to iterative set theory partly derives from the same philosophical picture.[30] Iterative set theory naturally takes objects, the ur-elements which are the members of sets, as primitive. They may be numbers or unspecified ur-elements, but they are naturally taken as given. Frege thought that an apriori discipline has to start from general principles. And it would be natural for him to ask where the ur-elements of set theory come from. If they were empirical objects, they would not be given apriori. He regarded the null set as an indefensible entity from the point of view of iterative set theory. It collects nothing. He thought a null entity (a null extension) is derivable only as the extension of an empty concept. If one took the numbers as primitive, one would not only be giving up logicism. One would be assuming particular objects without deriving their existence and character from general principles—thus controverting Frege's view of the nature of an apriori subject. If one could derive the existence of numbers from logical concepts, one would not need set theory to explain number theory or, Frege thought, for any other good purpose. Thus it would have been natural for him to see set theory as raising an epistemic puzzle about how its existence claims could be apriori, inasmuch as they appear to take statements about particulars as primitive or given.

Leibniz actually *characterizes* reason as the faculty for apprehending apriori, necessary truths. These include for him all mathematical truths. As I have noted, Leibniz regards all necessary truths as ultimately instances of, or derivative from, *general* logical principles together with definitional analyses and logical rules of inference.[31] Generality for Leibniz is a hallmark of human reason. Principles of identity and non-contradiction underlie and provide the logical basis for proof of mathematical truths. As noted, Frege agrees that arithmetic is thus derivative from general logical principles. He takes arithmetic to be an expression of pure reason, and its objects given directly to reason through logical principles (*FA* §105).

Kant famously separates apriority and necessity from pure reason in the sense that he holds that some apriori, necessary truths, the synthetic ones, can be known only by supplementing reason with the products of a non-rational faculty for producing singular representations—intuition. For Kant intuition is essentially a faculty for producing *singular* representations. It is part of his view that synthetic cognition of objects, including synthetic apriori cognition in arithmetic and geometry, must partly rest its justification on the deliver-

[30] *FA* §§ 46–54; *BL* 30; *GG* i. 2–3; *CP* 114, 227–228; *KS* 104–5, "Über formale Theorien der Arithmetik", O 96; *KS* 209–10, "Kritische Beleuchtung einiger Punkte in E. Schröders Vorlesungen über die Algebra der Logik", O 455. The latter passage especially seems to find the problem in the assumption of single things at the base of set theory. The idea that concepts are general and objects must be derivative from principles governing concepts guided his opposition.

[31] Leibniz, *New Essays*, Preface 49–50; I. i. 19, 83; IV. vii. 2–10, 408–13; IV. vii. 19, 424; IV. xi. 13, 445–6. On these matters, see Margaret Wilson, "Leibniz and Locke on 'First Truths' ", *Journal of History of Ideas*, 28 (1967), 347–66. Cf. n. 23.

ances of intuition. Hence the justification must rest partly on singular representations, and perhaps propositions or thoughts in singular form as well.

Of course, Frege disagrees with Kant about arithmetic. He holds that arithmetic is not synthetic, but analytic—at least in the sense that it is derivative from general logical principles without any need to appeal to intuition. But Frege purports to agree with Kant about geometry (*FA* § 89). He agrees that it is synthetic apriori. It is synthetic in that it is not derivable from logic. The logical coherence of non-Euclidean geometries seemed to confirm its synthetic character. Frege also purports to agree that geometry rests on pure apriori intuition.[32] He agrees with Kant in counting intuition a faculty different from the faculty of thought (e.g. *FA* §§ 26, 90). Frege's agreement with Kant that apriori truths of geometry rest on intuition, a faculty for producing singular representations, puts some pressure on Frege's view that apriori truths must rest on fundamentally general laws. As we shall see, there is some reason to think that Frege's relation to Kant on this matter is not as straightforwardly one of agreement as he represents it to be.

III

Let us now consider the second point of interest in Frege's characterization of apriority. This is his presumption that his characterization of apriority in terms of the primacy of generalizations in proof is equivalent with the usual post-Kantian characterization in terms of justificational independence from sense experience.

There are at least three areas where both the general characterization and Frege's assumption of equivalence can be challenged. One has to do with certain types of self-knowledge, and perhaps more broadly, certain context-dependent truths. One has to do with geometry. One has to do with arithmetic. I will not go into these issues in depth. But I hope that broaching them will be of both historical and substantive value.

Frege exhibits no interest in *cogito* judgements: judgements like the judgement that I am now thinking. But his characterization of apriority immediately rules them aposteriori, in view of the singularity of their form and their underivability from general laws. Now the question of whether *cogito* judgements are in fact aposteriori is a complex one.

Leibniz is in accord with Frege in counting them aposteriori. He counts them primitive, self-evident truths which nevertheless depend on 'experience'.[33] What Leibniz means by 'experience' is not very clear. His view suffers by

[32] Unlike Kant, Frege gives no clear evidence of believing that all synthetic apriori principles rest on intuition. He holds that the principle underlying (non-mathematical) induction is synthetic apriori, but he gives no reason to think that it rests on intuition. This point is made by Michael Dummett, "Frege and Kant on Geometry", *Inquiry*, 25 (1982), 240.

[33] Leibniz, *New Essays*, IV. vii. 7, 411; IV. ix. 3, 434; cf. IV. ii. 1, 367.

comparison to Kant's in its vastly less developed conception of cognitive faculties and of the nature of experience. Sometimes Leibniz associates experience with sense experience. But it appears that he sometimes uses a very broad conception of experience that would include any direct awareness of an object or event, whether or not this awareness proceeds through one of the senses. Thus 'experience' for Leibniz, at least at times, seems to include not only what we would count sense experience but intellectual 'experience' as well. A conception of apriority as independence from experience in this broad sense would be defensible. Its counting instances of the *cogito* 'aposteriori' would also be defensible.

Frege consistently associates experience with *sense experience*. If he were to relax this association, it would be open to him to side with Leibniz (or one side of Leibniz) here against Kant in counting non-sensory intellectual awareness of particular intellectual events as experience.[34] Such a conception would, however, sever the connection between apriority and independence of the experience of the senses. Frege seems to accept this connection. It has dominated conceptions of apriority since Kant. What seems to me thoroughly doubtful is that our cognition of instances of the *cogito* (and perhaps other indexical thoughts such as *I am here now* or *I exist*) is justificationally dependent on sense experience. Such cognition seems to depend only on intellectual understanding of the thought content in an instance of thinking it. Contingent, singular truths seem to be apriori in the sense that our warrant to accept them is justificationally independent of sense experience.

If these points are sound, they raise interesting questions about the relation between apriority, reason, and generality. It seems to me natural—at least as a working conjecture—still to regard reason (with Leibniz and Kant) as essentially involved in supplying general principles and rules of inference. A warrant can, however, be justificationally independent of sense experience if it gains its force from either reason or understanding. For Kant understanding essentially involves singular elements. For him reason is essentially general. Understanding, because of its interdependence with non-rational capacities, is sometimes understanding of truths in singular form that cannot be proved from general truths. Warrant can be apriori if it derives from reason or from understanding, if it does not depend on sense experience for any of the force of its epistemic warrant.

I believe that Kant was mistaken, however, in holding that understanding can yield non-logical cognition only if it applies to the form or deliverances

[34] Kant also thought that instances of the *cogito* produce no 'apriori' cognition. But this view cannot be directly derived from his characterizations of apriority and experience alone, as it can be from Leibniz's characterization. Rather Kant's view depends on his very complex (and I think mistaken) theory of the justificational dependence of cognition of one's thoughts in time on inner sense, which ultimately depends, albeit indirectly, for its justificational force on outer sense. I shall not discuss this Kantian view here.

of sensory capacities (and non-logical apriori cognition only if it applies to their form). I believe that understanding is capable of yielding non-empirical and non-sensible cognition of thoughts in singular form that are not derivable from general ones. One can, for example, know by intellection and understanding alone that certain of one's intellectual mental events are occurring (or have occurred), or that one is thinking. No invocation of sensible intuition or the form of one's sensory capacities is needed for the justification that underwrites the relevant knowledge. It seems to me plausible that our understanding sometimes applies to intentional contents that are tokens, instances of indexicals, in singular form.[35]

Perhaps to account for the apriority of our warrant for believing such instances, the warrant must be seen as deriving *partly or in some way* from something general. For example, to understand the self-evidence of an instance of *I am now thinking*, one must understand *I* according to the general rule that it applies to whomever is the author of the thought that contains its instantiation. One must understand a similar general rule for *now*. Thinking according to such rules, one can realize that any instance of *I am now thinking* will be true. This is an entirely general insight. It seems to me plausible to consider a logic for the *forms* of such indexicals as an expression for reason. Here the generality of reason would not reside in the form of the propositional content (which is singular), but in the generality of the rules governing its application. The semantical rule is in general form.[36]

But the realization of the truth of an *instance* of the *cogito* cannot be derived purely from these generalities. It cannot be derived purely from a logic of indexicals or from anything purely general. It must involve an awareness in understanding of an actual event of thinking and a recognition of its content. Thus the warrant cannot rest purely on an inference from general principles. There is something irreducibly singular in the application of the understanding. The warrant depends essentially for its force on the exercise of this singular application. Although the truth—the instance of the *cogito*—would count as aposteriori on Leibniz's conception and on Frege's conception, it is plausibly apriori on the Kantian conception: The warrant for an instance of thinking it is justificationally independent of sense experience. The warrant depends for its force purely on intellectual understanding applied to a singular instance of a *cogito* thought (cf. n. 34).

[35] Cf. Burge, "Our Entitlement to Self-Knowledge", *Proceedings of the Aristotelian Society*, 96 (1996), 91–116; *idem*, "Memory and Self-Knowledge", in P. Ludlow and N. Martin (eds.), *Externalism and Self-Knowledge* (Stanford, Calif.: CSLI Publications, 1998).

[36] For an example of a logic of such singular indexicals, see D. Kaplan, "A Logic of Demonstratives", in J. Almog, J. Perry, and H. Wettstein (eds.), *Themes from Kaplan* (New York: Oxford University Press, 1989).

IV

I turn now to Frege's application of his characterization of apriority to geometry. Frege accepted Kant's doctrine that Euclidean geometry is synthetic apriori. Frege meant by 'synthetic' here *not derivable from logic*. Frege also maintains with Kant that geometry rests on sensible geometrical spatial intuition. With Kant, Frege held the now discredited view that Euclidean geometry is both apriori and apriori-applicable to physical space. It is now tenable to hold that Euclidean geometry is apriori only if one considers it a pure mathematical discipline whose proper application, or applicability, to physical space is a separate and empirical question. I want, however, to discuss the issue of the epistemic status of Euclidean geometry from Frege's perspective.

What did Frege mean by his agreement with Kant about the epistemology of Euclidean geometry? There is no firm evidence that Frege accepted Kant's idealist conception of physical space. Frege's whole philosophy, especially in his mature period, seems out of sympathy with the explanation of apriority in terms of the mind's imposing its structure on the physical or mathematical worlds.[37] Frege articulated his agreement with Kant by agreeing that geometry is based on, or has its 'ground' in, pure intuition (*FA* §§ 12, 89).[38]

For Kant, pure intuition is both a faculty and one product of the faculty. Intuition is a faculty for singular, immediate representations. It represents singular elements of (or in) space or time without being mediated by any further representations that apply to the same semantical values or referents. Pure intuition is the faculty itself, considered independent of any passively received, sensational content. For Kant intuition could be either an intellectual faculty (in which case its exercises would always be pure), or a sensible one.[39] We humans have, according to Kant, only sensible intuition. Pure sensible intuition is the structure of the faculty which is constant regardless

[37] For an elaboration of some aspects of this theme, see Burge, "Frege on Knowing the Foundation" (Ch. 9 above).

[38] Cf. also "On a Geometrical Representation of Imaginary Forms in the Plane" in *CP* 1; *KS* 1, O 3–4; "Methods of Calculation based on an Extension of the Concept of Quantity" in *CP* 56; *KS* 50, "Rechnungsmethoden, die sich auf eine Erweiterung des Grössenbegriffes gründen", O 1.

[39] Frege shows a certain superficiality in his reading of Kant in *FA* § 12. There he first notes that in his *Logic* Kant defines an intuition as a singular representation, noting that there is no mention there of any connection with sensibility. He further notes that in the Transcendental-Aesthetic part of *Critique of Pure Reason* the connection is added (*hinzugedacht*), and must be added to serve as a principle of our cognition of synthetic apriori judgements. He concludes that the sense of the word 'intuition' is wider in the *Logic* than in the *Critique*. But it is not wider. In both books intuition is characterized in terms of singularity (and in the *Critique* sometimes in terms of immediacy as well). Cf. Kant, *Jäsche Logik* (1800), § I. 1; *Critique of Pure Reason*, A320/B376–7. Kant intentionally leaves sensibility out of the characterization of the notion in both books because he takes intellectual (non-sensible) intuition to be one possible type of intuition—possible in principle, though not for humans.

of what sensational contents one receives in sense-perceptual experience or produces in empirical imagination.

Kant also believed that pure sensible intuition could itself yield pure representations as product—pure formal intuitions.[40] Such representations are representations of elements in the structure of space and time. Given his idealism, these elements were supposed to be features of the structure of the faculty of sensible intuition. Intuitions of all sorts are characterized by Kant as being objective representations that are both immediate and singular.[41]

If one strips this view of its idealist elements, one can regard pure sensible intuition as a faculty for intuiting the pure structure (not of the faculty itself but) of mind-independent space and time. Frege shows no interest in pure temporal intuition. Of course, in his mature period he rejects Kant's view that arithmetic rests on pure temporal intuition, or intuition of any sort. He believed, however, that we have a capacity for pure spatial intuition. He believed that Euclidean geometry is in some way grounded in exercises of this capacity. Like Kant, Frege associates the capacity for pure intuition (in humans at least) with sensibility—the capacity for having sense experiences. He distinguishes it from a capacity for conceptual thought (*FA* §14).

What interests me is Frege's understanding of the singularity of pure intuition and its relation to his characterization of apriori truths as following from general principles that do not need or admit of proof. He cites and does not reject Kant's conception of intuitions as individual representations (*FA* §12). He regards the axioms and theorems of Euclidean geometry as apriori. So he thought that they are, or follow from, general principles that do not need or admit of proof. The proof must work out without reference to unprovable truths which are not general and which contain assertions about determinate objects (*bestimmte Gegendstände*). Kant takes intuitions to play a role in the warrant of some geometrical axioms and rules of inference. What is the epistemic role in Frege's view of pure intuitions—which for Kant are certainly singular, not general—in warranting the axioms of geometry?

Frege is aware of this question. He speaks to it in §13 of *The Foundations of Arithmetic*. He writes,

> One geometrical point, considered in itself, is not to be distinguished any way from any other; the same applies to lines and planes. Only if more points, lines, planes are comprehended at the same time in an intuition, does one distinguish them. From this it is explicable that in geometry general propositions are derived from intuition: the intuited points, lines, planes are really not particular (*besondern*) at all, and thus they can count as representatives of the whole of their kind. But with numbers it is different: each has its own particularity (*Eigentumlichkeit*).[42]

[40] Kant, *Critique of Pure Reason*, B160.

[41] Ibid. A320/B377.

[42] Frege does not make it clear why it matters that one can distinguish the objects of intuition from one another only if they are comprehended in a complex intuition, or why this fact shows that the objects are not really particular at all.

Frege does not use language in this passage that connects precisely with the language of his characterization of apriority.[43] Perhaps he simply believed that since the relevant objects of intuitions are not 'particular' (*besondern*), they are not 'determinate objects' (*bestimmte Gengenstände*). (Cf. the definition of 'aposteriority'.) Or perhaps he believed that pure intuition's contribution to the justification of general truths lies not in its representation of determinate objects (the individual lines and planes that it represents), but of aspects of them that are not particular to those objects. He may have thought that although we must be presented with particulars in pure intuition, the warranting power of the intuition lies *only* in geometrical properties that are invariant under Euclidean transformations. In either case, Frege does not give a precise explanation of how intuition helps 'ground' (*FA* § 12) our knowledge. Hence Frege gives no precise explanation of how his view of the apriority of geometry is compatible with his view of its depending on pure intuition—a faculty for singular representation.

Nevertheless, the main thrust of the passage seems to be to downgrade the role of the particularity of the geometrical objects, and of the singularity of thoughts about them, in the 'derivation' of general truths. In fact, Frege says that the objects of pure intuition in geometrical imagination are not genuinely particular. He seems to see the lines that he regards as objects of intuition as types. So they can serve as representatives whose characteristics that are shareable with relevantly similar objects are all that matter for arriving at general propositions. It is difficult to see here how Frege's view relates to Kant's, even bracketing the fact that Frege does not advocate Kantian idealism.

Let us approach this question by first comparing the just quoted passage from Frege with a passage in Leibniz. Leibniz writes:

> But I do not agree with what seems to be your view, that this kind of general certainty is provided in mathematics by 'particular demonstrations' concerning the diagram that has been drawn. You must understand that geometers do not derive their proofs from diagrams, although the expository approach makes it seem so. The cogency of the demonstration is independent of the diagram, whose only role is to make it easier to understand what is meant and to fix one's attention. It is universal propositions, i.e. definitions and axioms and theorems which have already been demonstrated, that make up the reasoning, and they would sustain it even if there were no diagram.[44]

Leibniz holds that the singular elements introduced through reliance on a diagram are inessential to a proof or derivation of the general propositions of

[43] In a paper on Hilbert, Frege seems to sympathize with the idea that geometrical axioms assert basic facts based on intuition. But he is focused on Hilbert's view that axioms both assert and define things. Frege's main point is that axioms cannot do both; he clearly believes that they assert something. There is little in the passage to help us with his attitudes towards the singularity of intuitions or their precise role in the epistemology of geometry. Cf. *CP* 275–7; *KS* 264–6, "Über die Grundlagen der Geometrie i" (1903), O 321–324.

[44] Leibniz, *New Essays*, IV. i. 360–1.

geometry. Frege's passage does not squarely advocate Leibniz's position. But Frege seems to be explaining away the elements of singularity in his conception of pure intuition in order to avoid acknowledging that the general truths of geometry are derivative in any way from singular elements in intuition. This direction of thought about (pure) geometry seems to me reasonable and plausible. But it is questionable whether Frege's view is really compatible with Kant's.

Kant sees himself as fundamentally at odds with Leibniz about geometry. He takes the role of pure intuition in geometry to be to produce an irreducibly singular element into mathematical understanding, reasoning, and justification. The problem for making these comparisons cleanly is that Kant's own view, though developed in great detail and subtlety, is not entirely clear or agreed upon.

I shall, however, sketch my view of it. Kant takes pure intuition in geometry to be intuitions of determinate objects. The objects of intuition are particulars, such as line-drawings, or even possible line-drawings, in pure geometrical intuition—pure imagination. (They can also be intuited in empirical intuition, on paper; but only non-empirical formal aspects of the empirical intuition play any role in mathematical understanding, reasoning, and justification.) From these objects one abstracts objects of a more general kind—'the triangle', for example—which are the objects of mathematical reasoning.[45] These latter objects are forms within the structure of space or time—on Kant's idealist view, forms of spatio-temporal intuition itself.

Theoretical cognition for Kant is fundamentally cognition of objects. Kant thought that pure mathematics has objects, and that those objects are not contingent, empirical objects.[46] 'Determination' (*Bestimmung*) is a fundamental term in Kant's epistemology. Objects of successful theoretical cognition—the sort yielded in geometry—are necessarily determinate, or objects of determinate concepts: specific, non-vague concepts. They are abstracted from determinate particulars that are referents of pure intuition. The abstracted objects are determinate formal objects—spatial shapes, like triangles, and lines, planes, volumes. They form the subject matter of Euclidean geometry. The principles of geometry are about these objects. And thoughts about them are supported and guided by pure intuition about particular instances of these determinate objects. The role of intuition, hence the role of representation of *particulars*, is ineliminable from Kant's account of our understanding and warrant for pure geometry.

[45] Kant, *Critique of Pure Reason*, A713–14/B741–2; A723/B751.

[46] The point is denied in M. Friedman, *Kant and the Exact Sciences* (Cambridge, Mass.: Harvard University Press, 1992), chs. 1 and 2. There are, however, numerous passages in which Kant makes it clear that he believes that pure mathematics has objects which are *not* the empirical objects experienced in space and time. For one such passage, see Kant, *Critique of Pure Reason*, A723/B751. I will develop these points in some detail in future work on Kant.

A passage in Kant that is comparable to the passages in Frege and Leibniz that we have just quoted is as follows:

Mathematical cognition [is reason-cognition out of] the construction of concepts. To construct a concept means to exhibit the intuition corresponding to it. For construction of a concept therefore a non-empirical intuition is required, which consequently as intuition is a single object (*einzelnes Objekt*), but nonetheless, as the construction of a concept (of a general representation), it [the intuition] must express in the representation general [or universal] validity (*Allgemeingültigkeit*) for all possible intuitions, which belong under the same concept. Thus I construct a triangle by exhibiting the object corresponding to this concept, either through mere imagination in pure intuition, or in accordance therewith also on paper through empirical intuition, but in both cases purely a priori, without having had to borrow the pattern for it from any experience. The single drawn figure is empirical, yet it serves to express the concept without impairing its universality (*Allgemeinheit*); for in the case of this empirical intuition we look only at the action of the construction of the concept, to which [concept] many determinations (*Bestimmungen*)—for example, the magnitude of the sides and angles—are completely indifferent, and therefore we abstract from these differences, which do not alter the concept of triangle.... mathematical cognition [considers] the general in the particular (*Besonderen*), in fact even in the individual (*Einzelnen*), although still a priori and by means of reason, so that just as this individual is determined under certain general conditions of construction, the object of the concept, to which this individual corresponds only as its schema, must be thought as universally (*allgemein*) determined.[47]

Frege's claim that 'the intuited points, lines, planes are really not particular (*besondern*) at all' is definitely not compatible with Kant's view. Kant maintains that the referents of intuition are always particular or singular.[48] He takes the singularity of the intuition to be essential to the normative, justificational account of mathematical cognition. He takes abstraction from certain particularities inherent in the single object presented in pure intuition (or even in empirical intuition) to be necessary to understanding the mathematical concept (the general concept, triangle) and to doing pure geometry. But the singularity of the intuition is irreducibly part of the justification of mathematical cognition.

Frege explains the general validity of geometrical truths by maintaining that the particularity of pure intuition is only apparent. They can therefore 'count as representatives of the whole of their kind'. Like Kant, he sees the particulars as serving as representatives or stand-ins for more general features. He does not explain what role the singular aspects of intuition play in the process. But unlike Kant, he appears to be committed to thinking that they

[47] Kant, *Critique of Pure Reason*, A713–14/B741–2. The translation 'we look at' and 'we abstract from' is necessary for smooth rendering in English, but the German uses an impersonal passive construction in both cases.

[48] Actually for Kant the immediate referents of intuitions are property instances or mark-instances had by particular objects. And objects include parts of space and time as well as physical objects. But these are subtleties that we need not go into here.

play no role in mathematical justification. This would explain his departure from Kantian doctrine in his claim that the intuited lines and so forth are not really particular at all. Unlike Kant, Frege is not interested in the particularity of mental acts in his explanation; this is a sign of his lack of commitment to Kantian idealism. He sees intuition as presenting typical geometrical structures which have no intrinsic individuality.

Kant explains the general validity of geometrical truths by maintaining that the particularity is genuine and ineliminable but is *used* as a schema. One abstracts from particular elements of the objects of intuition in forming a general object of the geometrical concept (and geometrical principle).

Like Frege, Kant does not make completely clear the role of the particular in warranting and guiding universal principles and inferential transitions. He seems to think that the particularistic elements in mathematical reasoning ground it in particular elements of space and time that reveal mathematical structures with maximum concreteness, and thus safeguard mathematical reasoning from the dangers that even transcendental philosophy is faced with. Kant seemed to think that mathematics' concern with particularity helps explain its certainty. But it *is* clear that he thought that the role of the particular is not to be explained away or seen as merely apparent. It is hard to escape the view that for Kant, in contrast to Frege, synthetic apriori propositions in geometry are grounded not in general propositions but in possible or actual particularistic judgements that are guided and supported by intuitions about particular, determinate objects of pure geometrical intuition. Although there are ways of understanding Frege's own view so as to render it internally consistent, and even perhaps sound, it is doubtful that it is consistent with Kant's.

Frege is aware of a need to discount the role of the particular, individual, or singular in geometrical warrant. If the general propositions rested, justificationally, on singular propositions, they could not be apriori in his sense.

Kant holds that the principles of geometry are strictly general or universally valid. He thinks that the basic principles are in the form of generalizations. But he does not hold that the root of geometrical warrant—the apriority of geometry—lies in generality. The synthetic apriori axioms—and the inferential transitions—in pure geometry rest on non-general representations, pure intuitions. His examples of pure intuition supplementing our conceptions to yield warranted belief commonly involve propositions used singularly about particular geometrical constructions in Euclidean space.[49]

Kant claims that the successive synthesis of the productive imagination in the generation of figures—a process of singular representation—is the basis of axioms and inferences in Euclidean geometry. Although the axioms are

[49] Kant, *Critique of Pure Reason*, A220–1/B267–8; A234/B287.

general, their warrant does not rest on general propositions or general thoughts alone.[50]

There is a way of construing Frege's introduction of the notion of apriority that would reconcile his view with Kant's. Recall that Frege writes: 'If . . . it is possible to derive the proof purely from general laws, which themselves neither need nor admit of proof, then the truth is a priori' (*FA* § 3). Geometrical proof, in the modern sense of 'proof', starts with geometrical axioms. These are general. Thus for Frege 'proof' in geometry rests on general truths, axioms. One might hold that Kant realized as well as anyone that geometrical proof begins with the axioms. On his own view, the axioms are general (universally quantified). Thus interpreted, there is no disagreement.

What makes this resolution unsatisfying to me is that neither Frege nor Kant utilized precisely this modern notion of proof. For Frege, proof is canonical justification. The axioms are, on his view, general, self-evident, and in need of no warrant from anything further. For Kant the axioms and proofs in geometry are warranted through their relations to actual or possible line-drawings in pure intuition—thus through their relation to singular representations.[51] These representations must (to represent their objects at all) be conceptualized and backed by propositions or judgements in singular form.

So, Frege's notion of proof is one of canonical justification, not merely deductive sequences of thoughts. And on Kant's view axioms and proofs in geometry require warrant from pure intuition, which is essentially a faculty of singular representation. Unlike Frege, Kant is not wedded to a view of apriority that takes it to be founded in generality.[52] For Kant, synthetic apriori cognition is cognition that is grounded in the particular. For Kant the use of pure intuition is an integral part of geometrical practice and the mathematical

[50] Of course, in his theory of arithmetic, Kant denies that arithmetical propositions are derivable from axioms—hence from anything general—at all. He seems to regard the singular arithmetical operations and equations as basic. Cf. ibid. A164–6/B204–6. Frege effectively criticized this extreme rejection of the role of axioms and proof in arithmetic (*FA* §5). He is of course right in rejecting Kant's view that intuition enters into the justification of *inferences* in geometry and arithmetic. The issue of whether particularity is basic to mathematical justification is independent of whether justification of mathematical propositions (commonly) involves proof, and even of whether particularity enters the justification through non-conceptual intuition or directly from understanding. For a fine discussion of Kant's view of the role of intuition in inferences, see Friedman, (*Kant and the Exact Sciences*, chs. 1 and 2). I believe that in supporting his sound view that Kant believed that intuition is necessary to mathematical inference, Friedman underplays the role of intuition in providing a basis for at least some of the axioms of Euclidean geometry. I think that Kant thought that intuitive constructions are as much a part of geometrical warrant and practice as commitment to the axioms is. Indeed the two go together.

[51] Ultimately for Kant the warrant presupposes the point that space is a form of our intuition of physical objects. Cf. Kant, *Critique of Pure Reason*, A46–8/B64–6; B147. Hence the warrant for geometry (and indeed all of mathematics) depends on the alleged fact that its applicability to the world of experience is guaranteed through its having as its subject matter the forms of our experience. This is part of Kant's 'transcendental deduction' of the objectivity of mathematics. I have little sympathy for this side of Kant's view, which in large part depends on his transcendental idealism.

[52] In fact, he contrasts apriori cognition in mathematics with apriori cognition in philosophy by insisting on the central role of particularity in the justification of mathematical cognition. Cf. ibid. A164/B204; A713–15/B741–3.

understanding of the axioms and inferences themselves. Thus insofar as it is possible to compare like to like—Frege's epistemological conception of proof with Kant's conception of justificational reasoning within geometry—the views of the two epistemologies appear quite different.

As I have emphasized, Frege leaves it unclear exactly what role intuition plays. But he implicitly denies a basic Kantian doctrine in holding that the objects of intuition are either not particular, or not fundamental to warrant in geometry. His picture of the role of particular elements in intuition seems in this respect to be more Leibnizian than Kantian. There is no evident room on his view to give intuition (as a singular representation) a warranting role.

I believe that Frege's verbal agreement with Kant about geometry is thus misleading. Frege accepts the language of Kant's doctrine of pure intuition—as applied to geometry. But it is doubtful that he can consistently accept all that Kant intends by this doctrine, and maintain the centrality of generality in his conception of apriority. Frege's Leibnizian conception of apriority takes generality of justificational starting point to be fundamental. He uses Kant's terminology of pure intuition, but he divests it of any commitment to referential singularity or reference to particulars, at least in its role in grounding geometrical principles. He retains Kant's view that intuition is essentially a non-rational (non-logical) faculty, thus appealing to intuition in order to explain his non-logicist, non-Leibnizian view of geometry. In this way he holds together a Leibnizian conception of apriority with a Kantian rejection of logicism about geometry. The fact that Frege provides a less detailed account of geometry, and less full explication of his term 'intuition' than Kant does, is explained by Frege's preoccupation with the mathematics of number.

There is a further aspect of Frege's account of intuition in geometry that renders it very different from Kant's. Kant takes intuition to be a type of *objective* representation.[53] Frege holds that intuition is not objective. In fact, he explains objectivity partly in terms of independence from intuition, which he regards as essentially subjective (*FA* § 26). In this passage, Frege makes his notorious claim that what is intuitable is not communicable. He sets out the thought experiment according to which what one being intuits as a plane another intuits as a point. He holds that since they can agree on geometrical principles (despite their subjective differences), their agreement is about something objective—about spatial structures that are subject to laws. Here again, it appears that particularistic aspects of intuition play no substantive role in Frege's account of the warrant for believing geometrical principles.

This doctrine of the subjectivity of spatial intuition is certainly not Kantian. Indeed, Kant characterizes intuition as an objective representation, in explicit contrast with subjective representations (sensations).[54] It is true

[53] Kant, *Critique of Pure Reason*, A320/B376–7.

[54] Ibid. A320/B377.

that from a transcendental point of view, Kant regards space itself and hence pure apriori intuition as a form of our 'subjective' constitution.[55] This is part of Kant's transcendental deduction of the objectivity of geometry. Kant thinks that only because, from the transcendental point of view, space, geometry, and apriori intuition are *all* to be construed idealistically as forms of the subject, can one account for the objectivity of apriori principles—and indeed the objectivity of pure intuition—in geometry about space. From the 'empirical point of view'—the point of view of the practice of ordinary science and mathematics—apriori intuition, geometrical principles, and space itself are all objectively valid and in no way confined to individuals' subjectivity.

Frege appears to have thought that the ability of mathematicians to produce logically coherent non-standard geometries shows that one can conceive (though not imagine or intuit) the falsity of Euclidean geometry. He thought that our grasp of the self-evidence of the axioms of Euclidean geometry depends on some non-rational, or at least non-logical, capacity that he termed 'intuition'. The elements intuited that are captured by the axioms are common to all—and in fact can be grasped in thought even by subjects whose subjective intuitions differ from ours.[56] So particularistic aspects of the intuitions seem to play no role in their warranting the axioms.

Frege calls Euclidean axioms self-evident. This view is in some tension with his appeal to intuition as grounds for the axioms. The warrant ('evidence') for believing the axioms seems not to rest purely in the senses of the axioms themselves. At least, one can apparently conceive of them as being false if one abstracts from spatial intuition. So the notion of self-evidence must be understood to include support from capacities whose deliverances are not entirely assimilated into the senses of the axioms themselves. Or at least such capacities provide a support for them that is needed as supplement to any conceptual grasp of them that would abstract from such support.[57] Perhaps

[55] Ibid. A48/B65.

[56] This explication is well expressed by Dummett, "Frege and Kant on Geometry", 250. I believe also that Dummett is correct in arguing that there is substantial evidence against the view that Frege accepted Kantian idealism about space. For an excellent, general discussion of Frege's views on geometry, see J. Tappenden, "Geometry and Generality in Frege's Philosophy of Arithmetic", *Synthese* 102 (1995), 319–61.

[57] It is not entirely clear to me what Frege, in his mature *post-Foundations* work, thought the relation between intuition and the senses of geometrical propositions is. The subjective elements in intuition are surely not part of the senses. Whether he thought that in conceiving non-Euclidean geometries and regarding them as logically consistent yet incompatible with Euclidean geometry, we give different senses to the key terms ('straight') or give the same sense but somehow abstract from intuitive support is not clear to me. Frege seems to have thought that sometimes intuitions are used in symbolic ways, as representations of something other than what is intuited, in geometrical reasoning. For example, in discussing generalizations of geometry beyond Euclidean space to a space of four dimensions, Frege says that intuition is not taken for what it is but as symbolic for something else (*FA* §14). He may have seen the same sort of process as involved in conceiving Euclidean geometry false in the context of reasoning within non-Euclidean geometry. This is a matter that invites further investigation.

general features associated with what mathematicians intuit, but *only* general features, play a role in warranting the axioms.

Both Kant and Frege held that Euclidean geometry yields apriori knowledge of physical space. As noted, this view is now untenable. What remains philosophically interesting is the epistemology of *pure* geometry. Warrant for mathematicians' belief in pure geometry seems to be apriori. Understanding the axioms seems sufficient to warrant believing them. But what does such understanding consist in? Geometrical concepts appear to depend in some way on a spatial ability. Although one can translate geometrical propositions into algebraic ones and produce equivalent models, the meaning of the geometrical propositions seems to me to be thereby lost. Pure geometry has some spatial content, even if it involves abstraction from the exact empirical structure of physical space. Perhaps there is something in common to all legitimate spatial notions that any pure geometry makes use of. Whether the role for a spatial ability in our warrant for believing them is particularistic and non-conceptual—as Kant claims—or fully general and conceptual—as Leibniz, and seemingly Frege, believe—seems to me to invite further investigation.

I believe that Kant is likely to be right about the dependence of our understanding of pure geometries on our representation of spatial properties through sensory, non-rational capacities. Frege appears to have sided with Kant on this matter. I think that Kant is probably wrong in holding that a non-conceptual capacity, pure intuition, plays a warranting role in geometrical understanding much less geometrical inference. Leibniz's view of warrant as deduction from basic (conceptually) understood truths of pure geometry seems closer to a sound modern mathematical epistemology. Like Kant, Frege appears to give pure intuition a role in warranting at least belief in the axioms of geometry. (I know of no evidence that Frege agreed with Kant that intuition is essential to warranting geometrical inference.) But Frege gives pure intuition a role in geometrical warrant only after removing the key Kantian feature of singularity of reference from this role. Moreover, Frege's view of the relation between the role of intuition in geometrical warrant and the alleged subjective character of intuition is left unclear.

It seems to me that conceptual understanding of the axioms of the various pure geometries suffices to warrant one in believing those axioms, as propositions in pure mathematics. Intuition in the Kantian sense seems to play a role in the fixing of geometrical content, but not in the warrant for believing the axioms or rules of inference.

V

I turn finally to the application of Frege's account of apriority to arithmetic. It is, of course, central to Frege's logicist project that truths about the numbers—which Frege certainly regarded as particular, determinate, formal objects (e.g. *FA* §§ 13, 18)—are derivative from general logical truths. The attempt to extract the existence and properties of particular objects from general principles centers, unfortunately, in Frege's defective Axiom V. There is a wide range of difficult issues here, and I cannot engage them seriously in this essay. But I want to broach, very briefly, some further points regarding Frege's characterization of apriority.

Suppose that Frege is mistaken, and arithmetic is not derivable in an epistemically fruitful way from purely general truths. Suppose that arithmetic has the form that it appears to have—a form that includes primitive singular intentional contents or propositions. For example, in the Peano axiomatization, arithmetic seems primitively to involve the thought that 0 is a number. And in normal arithmetical thinking we seem to know intentional contents that have singular form ($0 + 1 = 1$, for example) without deriving them from general ones. If some such knowledge is primitive—underived from general principles—then it counts as aposteriori on Frege's characterization. This would surely be a defect of the characterization. The knowledge does not seem to rest on anything other than arithmetic understanding. This seems to be intellectual understanding. The justification of the knowledge does not involve sense experience in any way. Even though the knowledge does not seem to rest on pure sensible intuition, or on anything having essentially to do with perceptual capacities, it may be irreducibly singular.[58] Indeed it seems to be irreducibly singular from an epistemic point of view, regardless of whether it concerns (as it appears to) abstract but particular objects. At any rate, the failure of Frege's logicism gives one reason to worry whether apriority and generality coincide, even in the case of arithmetic. It seems to me, even after a century of reductive attempts, that we need a deeper investigation into the epistemology of arithmetic.

I think that from an epistemological perspective, arithmetic should be distinguished from set theory, second-order logic, and various other parts of logic and mathematics. The enormous mathematical interest of the logicist project, and other reductive enterprises that have dominated the twentieth century, should not be allowed to obscure the fact that our understanding and hence our mode of knowing these other theories is different from our understanding of arithmetic. It seems to me even that the typical Peano formulation of arithmetic in terms of the successor function is epistemologically different from the formulation in terms of Arabic numerals on a base ten, which most of us learned first. Mathematical equivalence does not entail sameness of sense (in Frege's sense), and hence sameness of cognitive mode of presentation.

[58] In fact, our knowledge of set theory, while apriori, also seems to make primitive reference to particular sets, as noted earlier. Whereas Frege blamed set theory, rejecting it altogether, I am inclined to fault Frege's conception of apriority.

VI

It is time to summarize. Frege's characterization of apriority in terms of generality is a mischaracterization. Apriority bears an essential connection to justificational independence from experience. In modern times, 'experience' has come to mean *sense* experience. But Frege's characterization raises fundamental questions about the relation between apriority and generality. Frege followed a Leibnizian conception that assumed a close coincidence between the two notions.

If one thinks of experience sufficiently broadly (so as to include 'intellectual experience' not just sense experience), some of the pressure against the coincidence can be dissipated. Such a conception may have been one of Leibniz's conceptions of experience, and the associated conception of apriority may therefore have been Leibnizian as well. Such a conception could treat the instances of the *cogito* and other token, indexically based, self-evident truths as aposteriori. This is because the conception construes apriority in a way that excludes from the apriori even justificational dependence on purely intellectual 'experience'. Given a Kantian conception of apriority, which is more in line with the dominant modern conception, self-knowledge and knowledge of certain other indexical-involving truths can be apriori. For warrant seems to derive purely from intellectual understanding. It in no way rests on sense perception.

Problems with geometry and arithmetic remain. Leibniz, Kant, and Frege all maintained that geometry and arithmetic are apriori. If the position is carefully confined to pure geometry, it seems highly plausible. I believe, however, that we do not understand very well the role of spatial abilities in the content and justification of pure geometries. So I think that it is not fully clear whether justification in pure geometry rests on purely general propositions, although it seems to me likely that it does. The case of arithmetic is, I think, more serious as a possible counter-example to the claim of a coincidence between apriority and the primacy of generalizations in canonical justification. For arithmetic is apparently committed to basic truths in singular form, in its most natural and straightforward formulations.

I think that Frege is right to reject Kant's claim that the deliverances of a non-conceptual faculty, pure intuition, are *justificationally* basic in the warrant for arithmetic. But Kant may nevertheless have been right to hold that although cognition of arithmetic is apriori, cognition (or propositions) in singular form can be justificationally basic. One's justification derives from an understanding that encompasses singular intentional contents. On such a view, some apriority would be non-logical, and would not derive purely from *general* principles of pure reason. In arithmetic apriori knowledge would derive from intellectual, non-sense-perceptual understanding of necessary, non-context-dependent, singular intentional contents. I think that we should investigate in more depth the innovation that Kant offered: apriority that does not rest on logical or other general principles. I recommend doing so without assuming that apriori theoretical cognition must be constrained, as Kant insisted, by relation to sensibility. I recommend doing so without presuming that we must invoke Kant's notion of pure sensible intuition. I believe that we can follow Leibniz and Frege in avoiding essential reliance on pure intuition in arithmetic, without following them in insisting that generality lies at the base of all apriori warrant. Kant's conception of underived, singular understanding which is nevertheless apriori seems to be worth pursuing.

Postscript to "Frege on Apriority" (2003)

In both "Frege on Knowing the Foundation" (Chapter 9 above)[1] and "Frege on Apriority" (Chapter 10)[2] I note that on Frege's definitions of analyticity and apriority in section 3 of *Foundations of Arithmetic*, basic axioms of logic do not count as either analytic or synthetic. In both of these passages I express agreement with Michael Dummett that Frege's definitions constitute, in this respect, a harmless oversight. I still believe that the definitions' not covering these cases is harmless. I no longer believe that Frege was guilty of oversight.

Omission of basic logical axioms from the category *analyticity* goes back to Kant's own formulations. Kant did not count as analytic either the principle of non-contradiction or simple, strict identities like $a = a$.

Kant counted the principle of non-contradiction as a formal mark of truth and a principle of analytic cognition.[3] He seemed to regard such principles as fundamentally regulative meta-principles governing the practice of analysis. He did not count such principles themselves analytic.

Kant counted simple, strict identities tautologies. He states that tautologies are the limits of analysis.[4] He does not count the tautologies themselves analytic. Kant writes, "Analytic judgments are grounded in identity and can therein be resolved, but they are not identical for they need analysis and serve the explanation of concept."[5]

Kant's remark here explains the motivation for not counting basic principles or tautologies as analytic. They are not subject to analysis. Their truth is not revealed by analysis.

Of course, Kant does not count them synthetic either. The meta-principles are canons for regulation of good thinking. In their basic form, Kant may not have regarded them as true or false. The tautologies are certainly not synthetic since the content of their predicate does not "go beyond" the content of their subject.

Frege did not follow Leibniz and Kant in believing that the principle of non-contradiction and tautologies are all that lies at the foundation of logic. His logical principles are a much richer lot.

Frege also differed from Kant in his definition of analyticity. Kant's official characterization is in terms of containment: a truth or judgment is

[1] Cf. "Frege on Knowing the Foundation", Ch. 9, p. 322 and n. 6.
[2] "Frege on Apriority", Ch. 10, n. 18.
[3] Kant, *Critique of Pure Reason*, A59/B83–4, A151/B191. I am indebted to Verena Mayer for bringing to full conciousness an awareness of these points in Kant.
[4] Kant, *Gesammelte Schriften*, Königlich Preußischen Akademie der Wissenschaften, 29 vols. (Berlin: 1902–1983; 2nd edn. Berlin: De Gruyter, 1968, for vols. I–IX) *AA* IX, 111–12; *AA* XVI, 672–3.
[5] Ibid. *AA* XX, 322.

analytic if its predicate is contained in its subject.[6] Frege's is in terms of provability from logical laws: "If in [finding a proof of a proposition, and following it back to the primitive truths] we come only on general logical laws and on definitions, then the truth is an analytic one" (*Foundations of Arithmetic*, §3). Frege does associate provability with being contained in the axioms (*Begriffsschrift*, § 13). But containment is *explained* for Frege in terms of provability—most certainly not primarily in terms of conceptual containment.[7]

Frege's different conception of logic and his different axioms lead to his making different judgments about cases. For example, he takes statements of the form $a = a$ to be provable from general logical laws, hence both analytic and apriori.[8]

Both Kant's definition and Frege's, however, take analyticity to consist in being subject to analysis. In Kant's case, analysis is partitioning concepts. In Frege's case, it is finding a canonical proof from basic logical axioms. In each case, the starting points—the basic logical truths or principles are not subject to analysis. So they are not analytic. The motive for Kant's terminology carries over in different form in Frege's terminology. Both had a reason for not counting basic logical laws "analytic" in their own frameworks. Such laws are not subject to analysis—to further conceptual or proof-justificational unpacking.

Frege seems simply to have lined up his definition of apriority with his definition of analyticity. A truth's apriority lies in the nature of its proof. The basic axioms are self-evident. It is reasonable to believe them. They yield knowledge, and belief in them is surely warranted. But there is no discursive justification for them, no proof of them. I do not use either "analytic" or "apriori" in the way Frege did. But it is interesting to see that his use, at least of "analytic", derives from an old and well-motivated tradition.

6 Kant, *Critique of Pure Reason*, A6/B10. Evidently Kant did not take identity to be a type of containment.

7 One of Frege's great achievements in logic was to recognize that not all logical truths are truths of containment of one concept in another. Not all logical truths are in subject–predicate form.

8 Cf. the opening page of "On Sense and Denotation".

Bibliography

WORKS BY FREGE

In German

Begriffsschrift und andere Aufsätze (1879) ed. I. Angelelli (Hildesheim: Georg Olms, 1964; 2nd edn. 1977).
Die Grundgesetze der Arithmetik (1893, 1903; repr. Hildesheim: Georg Olms, 1962).
Die Grundlagen der Arithmetik (1884).
Gottlob Frege: Kleine Schriften, ed. I. Angelelli (Hildesheim: Georg Olms, 1967).
Logische Untersuchungen (Göttingen: Vandenhoeck und Ruprecht, 1966).
Nachgelassene Schriften, ed. H. Hermes, F. Kambartel, and F. Kaulbach (Hamburg: Felix Meiner, 1969; 2nd edn. 1983).
Wissenschaftliche Briefwechsel, ed. G. Gabriel, H. Hermes, F. Kambartel, C. Thiel, and A. Veraart (Hamburg: Felix Meiner, 1976).

Translations into English

"About the Law of Inertia", trans. R. Rand, *Synthese*, 13 (1961), 350–63.
The Basic Laws of Arithmetic: Exposition of the System, trans. and ed. M. Furth (Berkeley: University of California Press, 1967).
Begriffsschrift, in van Heijenoort ed., *Frege and Gödel* (Cambridge, Mass.: Harvard University Press, 1970).
Collected Papers on Mathematics, Logic, and Philosophy, ed. B. McGuinness (Oxford: Basil Blackwell, 1984).
Conceptual Notation, ed. T. Bynum (Oxford: Oxford University Press, 1972).
The Foundations of Arithmetic (1884), trans. J. L. Austin (Evanston, Ill.: Northwestern University Press, 1968; Oxford: Blackwell, 1980).
Logical Investigations, ed. P. Geach (New Haven: Yale University Press, 1977).
"On Sense and Nominatum", trans. H. Feigl, in H. Feigl and W. Sellars eds., *Readings in Philosophical Analysis* (New York: Appleton-Century-Crofts, 1949).
On the Foundations of Geometry and Formal Theories of Arithmetic, trans. and ed. E. W. Kluge (New Haven: Yale University Press, 1971).
Philosophical and Mathematical Correspondence, ed. B. McGuinness and H. Kaal (Chicago: University of Chicago Press, 1980).

Posthumous Writings, ed. H. Hermes, F. Kambartel, and F. Kaulbach, trans. P. Long and R. White (Chicago: University of Chicago Press, 1979).

Translations from the Philosophical Writings of Gottlob Frege, ed. P. Geach and M. Black, 2nd edn. (Oxford: Blackwell, 1966).

SECONDARY SOURCES

Works by Burge

"Reference and Proper Names", *The Journal of Philosophy*, 70 (1973), 425–39.

"Demonstrative Constructions, Reference, and Truth", *The Journal of Philosophy*, 71 (1974), 205–23.

"Belief *De Re*", *The Journal of Philosophy*, 74 (1977), 338–62.

"Kaplan, Quine, and Suspended Belief", *Philosophical Studies*, 31 (1977), 197–203.

"A Theory of Aggregates", *Nous*, 11 (1977), 97–117.

"Belief and Synonymy", *The Journal of Philosophy*, 75 (1978), 119–38.

"Self-Reference and Translation", in M. Guenthner-Reutter and F. Guenthner eds., *Meaning and Translation* (London: Duckworth, 1978), 137–53.

"Individualism and the Mental", *Midwest Studies of Philosophy*, 4 (1979), 73–121.

"Semantical Paradox", *The Journal of Philosophy*, 76 (1979), 169–98.

"Sinning Against Frege", *The Philosophical Review*, 88 (1979), 398–432; Ch. 5 this volume.

"Other Bodies", in A. Woodfield ed., *Thought and Object* (London: Oxford University Press, 1982).

"Russell's Problem and Intentional Identity", in James E. Tomberlin ed., *Agent, Language, and the Structure of the World* (Indianapolis: Hackett, 1983), 79–110.

"Frege on Extensions of Concepts, From 1884 to 1903", *The Philosophical Review*, 93 (1984), 3–34; Ch. 7 this volume.

Review of Michael Dummett, *The Interpretation of Frege's Philosophy, The Philosophical Review*, 93 (1984), 454–8.

"Frege on Truth", in L. Haaparanta and J. Hintikka eds., *Frege Synthesized* (Dordrecht: Reidel, 1986); Ch. 3 this volume.

"Intellectual Norms and Foundations of Mind", *The Journal of Philosophy*, 83 (1986), 697–720.

"Frege on Sense and Linguistic Meaning", in D. Bell and N. Cooper eds., *The Analytic Tradition* (Oxford: Blackwell, 1990); Ch. 6 this volume.

"Vision and Intentional Content", in E. Lepore and R. V. Gulick eds., *John Searle and his Critics* (Cambridge, Mass.: Blackwell, 1991).

"Frege on Knowing the Third Realm", *Mind*, 101 (1992), 633–49; Ch. 8 this volume.

"Philosophy of Language and Philosophy of Mind: 1950–1990", *The Philosophical Review*, 101 (1992), 3–51.

"Our Entitlement to Self-Knowledge", *Proceedings of the Aristotelian Society*, 96 (1996), 91–116.

"Frege on Knowing the Foundation", *Mind*, 107 (1998), 305–47; Ch. 9 this volume.

"Memory and Self-Knowledge", in P. Ludlow and N. Martin eds., *Externalism and Self-Knowledge* (Stanford, Calif.: CSLI Publications, 1998).

"Frege on Apriority", in P. Boghossian and C. Peacocke eds., *New Essays on the A Priori* (Oxford: Clarendon Press, 2000); Ch. 10 this volume.

"Logic and Analyticity", *Grazer Philosophische Studien*, 66 (2003), 199–249.

"Five Theses on *De Re* States and Attitudes", in a forthcoming book in honor of David Kaplan (Oxford: Oxford University Press).

Works by other authors

Ackermann, Diana, "Proper Names, Propositional Attitudes and Non-Descriptive Connotations", *Philosophical Studies*, 35 (1979), 55–69.

Adams, R. M., *Leibniz: Determinist, Theist, Idealist* (Oxford: Oxford University Press, 1994).

Barwise, Jon, and Perry, John, "Semantic Innocence and Uncompromising Situations", *Midwest Studies*, 6 (1981).

Bartlett, James M., *Funktion und Gegenstand* (Munich: M. Weiss, 1961).

Bays, Tim, "Tarski on Models", *The Journal of Symbolic Logic*, 66 (2001), 1700–26.

Benacerraf, Paul, "Mathematical Truth", *The Journal of Philosophy*, 70 (1973), 661–80; repr. in P. Benacerraf and H. Putnam eds., *Philosophy of Mathematics: Selected Readings* (Cambridge: Cambridge University Press, 1983).

—— "Frege: The Last Logicist", *Midwest Studies*, 6 (1981), 17–35.

Boisvert, Daniel R., and Lubbers, Christopher M., "Frege's Commitment to an Infinite Hierarchy of Senses", *Philosophical Papers*, 32 (2003), 31–64.

Boole, George, *Collected Logical Works*, ii (Chicago: Open Court Publishing Company, 1940).

Bynum, Terence Ward, "The Evolution of Frege's Logicism", in M. Schirn ed., *Studien zu Frege* (Stuttgart: Frommen-Holzboog, 1976).

Cantor, Georg, *Gesammelte Abhandlungen*, ed. E. Zermelo (Hildesheim: Georg Olms, 1962).

Carnap, Rudolf, *Meaning and Necessity* (1947), 2nd edn. (Chicago: University of Chicago Press, 1956, repr. 1967).

Castañeda, Hector-Neri, "On the Philosophical Foundations of the Theory of Communication: Reference", *Midwest Studies in Philosophy*, 2 (1977), 165–86.

—— "Perception, Belief, and the Structure of Physical Objects and Consciousness", *Synthese*, 35 (1977), 285–351.

Church, Alonzo, "The Calculi of Lambda-Conversion" (1941), in T. Burge *et al.* eds., *The Collected Works of Alonzo Church* (Cambridge, Mass.: MIT Press, forthcoming).

—— "Carnap's *Introduction to Semantics*", *The Philosophical Review*, 52 (1943), 298–305.

—— Review of Quine's "Notes on Existence and Necessity", *The Journal of Symbolic Logic*, 8 (1943), 45–7.

—— "A Formulation of the Logic of Sense and Denotation", in P. Henle, H. M. Kallen, and S. K. Langer eds., *Structure, Method, and Meaning* (New York: Liberal Arts Press, 1951).

—— *Introduction to Mathematical Logic*, i (Princeton: Princeton University Press, 1956).

—— "A Revised Formulation of the Logic of Sense and Denotation", *Nous*, 7 (1973), 24–33; (1974), 135–56.
—— "Schröder's Anticipation of the Theory of Types", *Erkenntnis*, 10 (1976), 407–11.
—— *The Collected Works of Alonzo Church*, ed. T. Burge *et al*. (Cambridge, Mass.: MIT Press, forthcoming).
Davidson, Donald, "Theories of Meaning and Learnable Languages" (1965), repr. in *Inquiries into Truth and Interpretation*, 2nd edn. (New York: Oxford University Press, 2001).
—— "On Saying That", *Synthese*, 19 (1969), 130–46.
—— "Reality without Reference", *Dialectica*, 31 (1977), 246–58.
—— *Truth and Predication* (Cambridge, Mass.: Harvard University Press, 2004).
Dedekind, Richard, *Gesammelte Mathematische Werke*, iii, ed. R. Fricke, E. Noether, and O. Ore (Braunschweig: Freidr. Vieweg & Sohn Akt.-Ges., 1932).
Donnellan, Keith, "Reference and Definite Descriptions", *The Philosophical Review*, 75 (1966), 281–304.
—— "Proper Names and Identifying Descriptions", *Synthese*, 21 (1970), 335–58; repr. in D. Davidson and G. Harman (eds.), *Semantics of Natural Language* (Dordrecht: Reidel, 1972).
Dummett, Michael, "Truth" (1959), repr. in *Truth and Other Enigmas* (Cambridge, Mass.: Harvard University Press, 1979).
—— "Gottlob Frege", in P. Edwards ed., *Encyclopedia of Philosophy* (New York: Macmillan Co. and the Free Press, 1967).
—— *Frege: Philosophy of Language* (London: Duckworth; New York: Harper and Row, 1973); 2nd edn. (Cambridge, Mass.: Harvard University Press, 1981).
—— *Truth and Other Enigmas* (Cambridge, Mass.: Harvard University Press, 1978).
—— "Indexicality and *Oratio Obliqua*", in *The Interpretation of Frege's Philosophy* (Cambridge, Mass.: Harvard University Press, 1981).
—— *The Interpretation of Frege's Philosophy* (Cambridge, Mass.: Harvard University Press, 1981).
—— "Frege and Kant on Geometry", *Inquiry*, 25 (1982), 233–54.
—— *Frege: Philosophy of Mathematics* (Cambridge, Mass.: Harvard University Press, 1991).
—— *Origins of Analytic Philosophy* (Cambridge, Mass.: Harvard University Press, 1994).
Evans, Gareth, *The Varieties of Reference* (Oxford: Clarendon Press, 1982).
Friedman, Michael, *Kant and the Exact Sciences* (Cambridge, Mass.: Harvard University Press, 1992).
Furth, Montgomery, "Two Types of Denotation", in *Studies in Logical Theory*, American Philosophical Quarterly Monograph Series, 2 (Oxford: Blackwell, 1968).
Gabriel, Gottfried, "Frege's 'Epistemology in Disguise' ", in M. Schirn ed., *Frege: Importance and Legacy* (Berlin: Walter de Gruyter, 1996).
Gödel, Kurt, "Russell's Mathematical Logic" (1944), in P. Benacerraf and H. Putnam eds., *Philosophy of Mathematics* (Englewood Cliffs, NJ: Prentice-Hall, 1964; Cambridge: Cambridge University Press, 1983).
Hamilton, Sir William, *Lectures on Logic*, ii (New York: Sheldon and Company, 1858).

Heck, Richard, "Frege and Semantics", in T. G. Ricketts ed., *The Cambridge Companion to Frege* (Cambridge: Cambridge University Press, forthcoming).

Jevons, W. Stanley, *The Principles of Science* (1865) (London: Macmillan and Co., 1920).

Kant, Immanuel, *Critique of Pure Reason* (1781, 1787).

——*Critique of Judgement* (1790).

——*Jäsche Logik* (1800).

——*Gesammelte Schriften*, Königlich Preußische Akademie der Wissenschaften, 29 vols. (Berlin: 1902–1983; 2nd edn., Berlin: De Guyter, 1968, for vols. I–IX).

Kaplan, David, "Foundations of Intensional Logic" (dissertation, UCLA, 1964).

——"Quantifying In", in D. Davidson and J. Hintikka eds., *Words and Objections* (Dordrecht: Reidel, 1969).

——"Demonstratives" (circulated mimeograph, 1977).

——"A Logic of Demonstratives", in J. Almog, J. Perry, and H. Wettstein eds., *Themes from Kaplan* (New York: Oxford University Press, 1989).

Klemke, E. D. (ed.), *Essays on Frege* (Urbana, Ill.: University of Illinois Press, 1968).

Kripke, Saul, "Naming and Necessity", in D. Davidson and G. Harman eds., *Semantics of Natural Language* (Dordrecht: Reidel, 1972).

——*Naming and Necessity* (Cambridge, Mass.: Harvard University Press, 1980).

Lakatos, Imre, *Proofs and Refutations* (Cambridge: Cambridge University Press, 1976).

Leibniz, Gottfried W., *Discourse on Metaphysics* (1686) in *Philosophical Essays*, ed. R. Ariew and D. Garber (Indianapolis: Hackett Publishing Co., 1989).

——"Monadology" (1714), in *Philosophical Essays*, ed. R. Ariew and D. Garber (Indianapolis: Hackett Publishing Co., 1989).

——*New Essays on Human Understanding* (1765), trans. P. Remnant and J. Bennett (Cambridge and New York: Cambridge University Press, 1989).

——*Leibniz: Logical Papers*, trans. and ed. G. H. R. Parkinson (Oxford: Clarendon Press, 1966).

——*Philosophical Essays*, ed. R. Ariew and D. Garber (Indianapolis: Hackett Publishing Co., 1989).

Lewis, David K., "Anselm and Actuality", *Nous*, 4 (1970), 175–88.

Linsky, Leonard, *Referring* (New York: Humanities Press, 1967).

Loar, Brian, "The Semantics of Singular Terms", *Philosophical Studies*, 30 (1976), 353–77.

Lotze, R. H., *Logik* (Leipzig, 1880).

——*Logic*, trans. B. Bosanquet (Oxford, 1888; repr. New York, 1980).

Marcus, Ruth, Review of Leonard Linsky, *Names and Descriptions, The Philosophical Review*, 87 (1978), 497–504.

Mates, Benson, "Synonymity", in L. Linsky ed., *Semantics and the Philosophy of Language* (Urbana, III.: University of Illinois Press, 1952).

Meyer, Albert R., "What is a Model of the Lambda Calculus?" (xerox, Laboratory for Computer Science, MIT, Cambridge, Mass.).

Mill, J. S., *A System of Logic* (1843) (London: Longmans, Green and Co., 1865; New York: Harper & Bros., 1893).

Parsons, Charles, "Frege's Theory of Number", in M. Black ed., *Philosophy in America* (Ithaca, NY: Cornell University Press, 1965).

—— "Sets and Classes", *Nous*, 8 (1974), 1–12.
—— "Some Remarks on Frege's Conception of Extension", in M. Schirn ed., *Studien zu Fregen* (Stuttgart: Frommen-Holzboog, 1976).
—— "What is the Iterative Conception of Set?", in R. E. Butts and J. Hintikka eds., *Logic, Foundations of Mathematics, and Compatibility Theory* (Dordrecht: Reidel, 1977).
Parsons, Terence, "Frege's Hierarchies of Indirect Sense and the Paradox of Analysis", *Midwest Studies in Philosophy*, 6 (1981), 37–57.
—— "Why Frege Should Not Have Said 'The Concept Horse is not a Concept' ", *History of Philosophical Quarterly*, 3 (1986), 449–65.
—— "A Quasi-Fregean-Carnapian-Early Kaplanian Semantics", in a volume honoring David Kaplan (Oxford: Oxford University Press, forthcoming).
Peacocke, Christopher, "Entitlement, Self-Knowledge, and Conceptual Redeployment", *Proceedings of the Aristotelian Society*, 96 (1996), 117–58.
—— *Being and Being Known* (Oxford: Oxford University Press, 1999).
Perry, John, "Frege on Demonstratives", *The Philosophical Review*, 86 (1977), 474–97.
Plantinga, Alvin, "The Boethian Compromise", *American Philosophical Quarterly*, 15 (1978), 129–38.
Putnam, Hilary, "Is Semantics Possible?" (1970), in *Philosophical Papers*, ii (Cambridge: Cambridge University Press).
—— "The Meaning of 'Meaning' " (1975), in *Philosophical Papers*, ii (Cambridge: Cambridge University Press, 1975).
—— *Reason, Truth, and History* (Cambridge: Cambridge University Press, 1981).
Quine, W. V., "Two Dogmas of Empiricism" (1953), repr. in *From a Logical Point of View* (New York: Harper and Row, 1961).
—— "Carnap and Logical Truth" (1954), in *Ways of Paradox* (New York: Random House, 1966).
—— *Word and Object* (Cambridge, Mass.: MIT Press, 1960).
—— *From a Logical Point of View*, 2nd edn. (New York: Harper Torchbooks, 1961).
—— *Ontological Relativity* (New York: Columbia University Press, 1969).
Ricketts, Thomas G., "Objectivity and Objecthood: Frege's Metaphysics of Judgement", in *Frege Synthesized* (Dordrecht: Reidel, 1986).
—— "Logic and Truth in Frege", *Proceedings of the Aristotelian Society*, suppl. vol. 70 (1996), 121–40.
Russell, Bertrand, *The Principles of Mathematics* (New York: W. W. Norton, 1902).
—— "On Denoting", *Mind*, 14 (1905), 479–93.
—— "The Philosophy of Logical Atomism" (1918), in R. C. Marsh ed., *Logic and Knowledge: Essays 1901–1950* (London: Unwin Hyman, 1956).
—— and Whitehead, A. N., *Principia Mathematica*, i (New York: Cambridge University Press, 1910).
Schirn, M. ed., *Studien zur Frege* (Stuttgart: Frommen-Holzborg, 1976).
—— ed., *Frege: Importance and Legacy* (Berlin: Walter de Gruyter, 1996).
Schröder, Ernst, *Vorlesungen über die Algebra der Logik*, i (Bronx, NY: Chelsea Publishing Company, 1966).
Scott, D. S., "Lambda Calculus: Some Models, Some Philosophy", in J. Barwise, H. J. Keisler, and K. Kunen eds., *The Kleene Symposium* (Amsterdam: North-Holland, 1980).

Searle, John, "Proper Names", *Mind*, 67 (1958), 166–73; repr. in J. Rosenberg and C. Travis eds., *Readings in the Philosophy of Language* (Englewood Cliffs, NJ: Prentice-Hall, 1971).

Sluga, Hans D., "Frege and the Rise of Analytic Philosophy", *Inquiry*, 18 (1975), 471–98.

—— *Gottlob Frege* (London: Routledge and Kegan Paul, 1980).

Soames, Scott, "The Modal Argument: Wide Scope and Rigidified Descriptions", *Nous*, 32 (1998), 1–22.

Stanley, Jason, "Truth and Meta-Theory in Frege", *Pacific Philosophical Quarterly*, 77 (1996), 45–70.

Strawson, Peter, *Individuals* (1959) (Garden City, NY: Anchor, 1963).

Tappenden, Jamie, "Geometry and Generality in Frege's Philosophy of Arithmetic", *Synthese*, 102 (1995), 319–61.

Tarski, Alfred, "On the Concept of Logical Consequence" (1936), in *Logic, Semantics, Metamathematics*, 2nd edn. (Indianapolis: Hackett, 1983).

—— "The Concept of Truth in Formalized Languages, in *Logic, Semantics, Metamathematics* (Cambridge: Cambridge University Press, 1956).

Veroort, Albert, "Geschichte des wissenschaftlichen Nachlasses Gottlob Freges und seiner Edition. Mit einem Katalog des ursprunglichen Bestands der nachgelassene-Schriften Freges", in M. Schirn ed., *Studien zu Frege* (Stuttgart: Frommen-Holzborg, 1976), i, item 47, p. 95.

Wallace, John, *Philosophical Grammar* (dissertation, Stanford, 1970).

—— "Belief and Satisfaction", *Nous*, 6 (1972), 85–95.

—— "Only in the Context of a Sentence do Words have Meaning", in *Midwest Studies* 2 (1977).

Weiner, Joan, *Frege in Perspective* (Ithaca, NY: Cornell University Press, 1990).

Wilson, Margaret, "Leibniz and Locke on 'First Truths' ", *Journal of the History of Ideas*, 28 (1967), 347–66.

Wittgenstein, Ludwig, *Philosophical Investigations*, trans. G. E. C. Anscombe (Oxford: Blackwell, 1953).

Zermelo, Ernst, "A New Proof of the Possibility of a Well-Ordering" (1908), in J. van Heijenoort ed., *From Frege to Gödel* (Cambridge, Mass.: Harvard University Press, 1981).

Author Index

Citation Index

FREGE'S PUBLICATIONS

FREGE'S POSTHUMOUSLY PUBLISHED WORK

LETTERS

Subject Index

The following index is meant to aid serious study. The main objective of the index is not to provide a quick look-up, or a browser's aid. It is to provide a way to reconstruct the paths of all the major themes through the volume. Entries contain numerals marking all, or almost all, discussions of the topic marked by the entry. The reader who wants to find *basic* or *elementary* passages for a given topic will find such passages marked by *italicized* numerals. These occur in all the larger entries which constitute distinctively philosophical vocabulary, or whose usage derives from Frege's philosophy or my interpretation of it.

The cross-references are arranged in a tree structure. Broadly speaking, the more general and more explanatorily basic entries contain cross-references (marked by "*see also*") to the more specific and less explanatorily basic entries. Occasionally, the more general entry is less explanatorily basic, in which cases, I have made a choice. This method of ordering also applies within the cross-references of a given entry. Thus entries for subject matters ("numbers") come before sciences about the subject matter ("arithmetic") or other representations of the subject matter ("knowledge, mathematical"). Ordinary or generic representation ("perception", "thought", "knowledge") comes before reflection on it ("epistemology"). The scheme inevitably raises philosophical questions. I have tried to provide an ordering that corresponds to Frege's conception of explanatory order, as far as possible, within the confines of the structure. Thus, in the tree structure, "truth" comes before "thought" ("thought content"); and "thought" comes before "language" and before "propositional attitudes". Fregean views or terminology come before non-Fregean views or contrasting terminology.

There are many limitations to any such scheme. Some notions are involved in mutual or holistic explanation. There are different types of explanation. One type may serve one order; another type may underwrite a different order. Some notions are listed for contrast, not explanation. Still, the index may be used to stimulate reflection on the explanatory and logical structures of Frege's philosophy, as well as to help find passages. Happy hunting!—Tyler Burge