

# WHAT DOES IT MEAN TO SAY THAT LOGIC IS FORMAL?

by

John Gordon MacFarlane

A.B., Philosophy, Harvard College, 1991

M.A., Philosophy, University of Pittsburgh, 1994

M.A., Classics, University of Pittsburgh, 1997

Submitted to the Graduate Faculty of  
Arts and Sciences in partial fulfillment  
of the requirements for the degree of  
Doctor of Philosophy

University of Pittsburgh

2000

---

Robert Brandom, Distinguished Service Professor of Philosophy  
(Director)

---

Nuel Belnap, Alan Ross Anderson Distinguished  
Professor of Philosophy (Second Reader)

---

Joseph Camp, Professor of Philosophy

---

Danielle Macbeth, Associate Professor of Philosophy,  
Haverford College (Outside Reader)

---

Kenneth Manders, Associate Professor of Philosophy

---

Gerald Massey, Distinguished Service Professor of Philosophy

## WHAT DOES IT MEAN TO SAY THAT LOGIC IS FORMAL?

John Gordon MacFarlane, PhD

University of Pittsburgh, 2000

Much philosophy of logic is shaped, explicitly or implicitly, by the thought that logic is distinctively *formal* and abstracts from material content. The distinction between formal and material does not appear to coincide with the more familiar contrasts between *a priori* and empirical, necessary and contingent, analytic and synthetic—indeed, it is often invoked to *explain* these. Nor, it turns out, can it be explained by appeal to schematic inference patterns, syntactic rules, or grammar. What does it mean, then, to say that logic is distinctively formal?

Three things: logic is said to be formal (or “topic-neutral”)

- (1) in the sense that it provides constitutive norms for thought as such,
- (2) in the sense that it is indifferent to the particular identities of objects, and
- (3) in the sense that it abstracts entirely from the *semantic content* of thought.

Though these three notions of formality are by no means equivalent, they are frequently run together. The reason, I argue, is that modern talk of the formality of logic has its source in Kant, and these three notions come together in the context of Kant’s transcendental philosophy. Outside of this context (e.g., in Frege), they can come apart. Attending to this history can help us to see the sources of our disparate intuitions about logic, and more importantly to sort these intuitions into central and adventitious ones. I argue that we have largely lost sight of the notion of formality (1) by which logic was demarcated in a central tradition from Leibniz through Frege—the intellectual home of most of the philosophical projects for which it matters how logic is demarcated.

This historical perspective is especially useful in evaluating contemporary debates about the demarcation of logic, which often seem to turn on opposing but equally brute intuitions

about logicality. As an illustration, I examine the popular permutation-invariance account of logicality, which is commonly motivated by appeal to sense (2) of formality. I present the account in a way that reveals a hidden lacuna, and I show how this lacuna might be filled by appealing to formality in sense (1).

# PREFACE

The germ of this project was James Allen’s interesting seminar on the development of Aristotle’s logic. In what sense, I wondered, can the hodge-podge of argumentative advice and quasi-logical rules contained in the *Topics* be called a “logic”? This question led me to the contemporary literature on the demarcation of logic. One of the things that struck me immediately was the sheer variety of proposals. I wondered whether they were all aimed at the same target.

I worked through some of this literature in an independent study with John McDowell. One of the books I read was John Etchemendy’s *The Concept of Logical Consequence*. Etchemendy argues that the hunt for a principled dividing line between logical and non-logical constants is a misguided by-product of Tarski’s account of logical consequence. This struck me as wrong. Surely the demarcational enterprise has deeper roots than that: even the medieval logicians distinguished between “formal” and “material” consequence. And so I began to investigate the history of conceptions of logicity, from Aristotle to the present. It soon became clear that there was an interesting story to tell about the persistent idea that logic is distinctively “formal.”

There are many people to thank. Bob Brandom was an ideal dissertation director; his unflagging enthusiasm for the project and probing comments on my work in progress were invaluable. Danielle Macbeth went far beyond what is usually expected of an outside reader; I learned much from reading and reacting to drafts of her book on Frege’s logic and from our e-mail discussions of Kant and Frege. I thank Nuel Belnap for his detailed comments (especially on chapter 6) and for warning me away from many confusions, Joe Camp for

asking a crucial question that led me investigate the source of Kant's logical hylomorphism, Ken Manders for referring me to a key passage in Frege's "On the Foundations of Geometry: Second Series," Jerry Massey for pointing me to Church's review of Carnap's *Formalization of Logic* and to his own work in the philosophy of logic, Nicholas Rescher for leading me to Trendelenburg and Couturat, Steve Engstrom for reading and discussing chapter 4, Rega Wood for comments on appendix A, James Conant for conversations about Kant, Russell, and Wittgenstein, and Lionel Shapiro and John Roberts for listening to my half-baked ideas and responding with encouragement and helpful comments. I also profited from the comments of many people who responded to talk versions of some of this material, including Joseph Almog, Lanier Anderson, Nick Asher, Paul Benacerraf, Ned Block, Tyler Burge, John Burgess, John Carriero, Charles Chihara, John Etchemendy, Hartry Field, Kit Fine, Peter Godfrey-Smith, Gil Harman, Mark Johnston, David Kaplan, Paolo Mancosu, Tony Martin, Calvin Normore, Gideon Rosen, Stephen Schiffer, Hans Sluga, David Sosa, Peter Unger, and Dan Warren. Finally, I thank Colleen Boyle for keeping my spirits up throughout.

This dissertation is dedicated to the memory of my grandfathers, John C. MacFarlane and Gordon A. Wilson. I like to think that the very different lines along which they inspired me intersect in this document.

# Contents

<b>1</b>	<b>INTRODUCTION</b>	<b>1</b>
1.1	What is logic? . . . . .	1
1.1.1	Method, not subject matter? . . . . .	2
1.1.2	Enthymemes . . . . .	3
1.1.3	Analyticity, a prioricity, necessity . . . . .	4
1.1.4	Formal vs. material . . . . .	6
1.2	Why do we need a demarcation? . . . . .	7
1.2.1	Logicism . . . . .	7
1.2.2	Structuralism . . . . .	11
1.2.3	Explicitation . . . . .	14
1.3	Why formality? . . . . .	15
1.3.1	Pragmatic vs. principled demarcations . . . . .	16
1.3.2	Alternatives . . . . .	19
1.3.3	Historical importance of formality . . . . .	20
1.4	Why a historical approach? . . . . .	22
1.5	Prospectus . . . . .	27
<b>2</b>	<b>DECOYS</b>	<b>31</b>
2.1	Syntactic formality . . . . .	32
2.2	Schematic formality . . . . .	36
2.3	Grammatical formality . . . . .	41

2.3.1	The immanence of grammar . . . . .	42
2.3.2	Identity . . . . .	44
2.3.3	Grammatical chauvinism . . . . .	46
2.4	Conclusion . . . . .	49
<b>3</b>	<b>THREE NOTIONS OF LOGICAL FORMALITY</b>	<b>50</b>
3.1	1-formality . . . . .	52
3.2	2-formality . . . . .	56
3.3	3-formality . . . . .	60
3.3.1	Schlick and Einstein . . . . .	62
3.3.2	Carnap . . . . .	63
3.3.3	Nagel . . . . .	65
3.4	Independence of the three notions . . . . .	65
3.5	Formality, generality, and topic-neutrality . . . . .	69
3.6	Philosophical significance of the three notions . . . . .	76
3.7	A puzzle . . . . .	77
<b>4</b>	<b>KANT AND THE FORMALITY OF LOGIC</b>	<b>79</b>
4.1	Kant's characterization of logic . . . . .	81
4.1.1	Kant's taxonomy of logics . . . . .	81
4.1.2	What does Kant mean by "general"? . . . . .	84
4.1.3	What does Kant mean by "formal"? . . . . .	87
4.1.4	Generality, formality, and Kant's demarcation of logic . . . . .	91
4.2	The originality of Kant's characterization . . . . .	95
4.2.1	The Wolffian school . . . . .	96
4.2.2	Descartes and Locke . . . . .	99
4.2.3	Leibniz . . . . .	102
4.2.4	Scholastics . . . . .	110
4.2.5	Crusius, Lambert, Tetens . . . . .	113

4.3	The genesis of logical hylomorphism in Kant . . . . .	114
4.3.1	The definition of logic . . . . .	114
4.3.2	Material criterion of truth . . . . .	115
4.3.3	Remarks on the history of logic . . . . .	117
4.3.4	Evidence of the Reflexionen . . . . .	118
4.3.5	Precritical works . . . . .	119
4.4	Transcendental idealism and the formality of logic . . . . .	121
4.5	Kant as the source of modern logical hylomorphism . . . . .	127
4.5.1	In Germany . . . . .	127
4.5.2	In Britain . . . . .	128
4.6	Conclusion . . . . .	133
<b>5</b>	<b>FREGE AND THE FORMALITY OF LOGIC</b>	<b>135</b>
5.1	The status of logic in Frege's early works . . . . .	137
5.2	Frege's rejection of Kant's Thesis . . . . .	143
5.2.1	Frege and Kant on logic and arithmetic . . . . .	143
5.2.2	Frege's characterization of logic as 1-formal . . . . .	147
5.2.3	"The legend of the sterility of pure logic" . . . . .	149
5.2.4	"Logical form" in the FA . . . . .	151
5.2.5	Frege's rejection of 3-formality . . . . .	152
5.2.6	"The most general laws of truth" . . . . .	157
5.2.7	How Frege resists the Kantian argument . . . . .	161
5.3	Is Frege's logic 2-formal? . . . . .	165
5.4	Does Russell's Paradox vindicate Kant's Thesis? . . . . .	167
5.5	Frege's influence . . . . .	170
<b>6</b>	<b>PERMUTATION INVARIANCE AND LOGICALITY</b>	<b>173</b>
6.1	Permutation invariance . . . . .	177
6.1.1	Categorial grammar . . . . .	178

6.1.2	Typed presemantics . . . . .	180
6.1.3	Compositional semantics . . . . .	181
6.1.4	The permutation invariance criterion . . . . .	182
6.1.5	Advantages of this framework . . . . .	185
6.2	Logical notions and logical constants . . . . .	187
6.2.1	Presemantics and semantics . . . . .	187
6.2.2	Logical notions and logical constants . . . . .	188
6.2.3	Quantifiers and domains . . . . .	193
6.2.4	A note on the status of $\lambda$ . . . . .	195
6.3	Invariance over variable domains . . . . .	196
6.3.1	What does the quantifier domain represent? . . . . .	196
6.3.2	Quantifier domains as indexical types . . . . .	199
6.3.3	Semantics with indexical types . . . . .	201
6.3.4	The invariance criterion restated . . . . .	203
6.3.5	Logical constants revisited . . . . .	204
6.3.6	Advantages of this approach . . . . .	205
6.4	Extending the account to sentential operators . . . . .	207
6.5	Extending the account to intensional logics . . . . .	214
6.6	Limitations of the invariance criterion . . . . .	219
6.7	Intrinsic structure: a proposal . . . . .	223
6.7.1	Two sources of constraints on presemantic types . . . . .	224
6.7.2	What is intrinsic structure? . . . . .	227
6.7.3	Application: multivalued logics . . . . .	229
6.7.4	Application: modal logics . . . . .	231
6.7.5	Application: tense logic . . . . .	232
6.7.6	Application: extensional logic . . . . .	234
6.8	Conclusion . . . . .	235

<b>7</b>	<b>CONCLUSION</b>	<b>238</b>
7.1	After formality? . . . . .	238
7.2	How we got where we are . . . . .	240
7.3	Applications . . . . .	242
7.3.1	Topic-neutrality . . . . .	242
7.3.2	Permutation invariance . . . . .	243
7.3.3	The debate over second-order logic . . . . .	244
7.3.4	The logicity of Hume’s Principle . . . . .	246
7.4	The centrality of 1-formality . . . . .	249
7.5	Methodological postscript . . . . .	252
<b>A</b>	<b>THE ORIGINS OF LOGICAL HYLOMORPHISM</b>	<b>255</b>
A.1	Aristotle and the commentary tradition . . . . .	258
A.1.1	Aristotle and formal logic . . . . .	258
A.1.2	Logical hylomorphism in the Greek commentaries . . . . .	260
A.1.3	Logical hylomorphism in Boethius . . . . .	264
A.2	Unmethodically conclusive arguments . . . . .	266
A.2.1	The debate over unmethodically conclusive arguments . . . . .	268
A.2.2	The neoterioi’s case . . . . .	269
A.2.3	The Peripatetic response . . . . .	270
A.2.4	Galen’s pragmatic alternative . . . . .	272
A.3	The transformation of the topics . . . . .	274
A.3.1	Aristotle’s Topics . . . . .	274
A.3.2	Topics as axioms . . . . .	275
A.3.3	Early medieval theories of Topics . . . . .	278
A.4	Abelard on perfect and imperfect inferentia . . . . .	281
A.4.1	Topics and the grounds of inference . . . . .	282
A.4.2	Generality and abstraction from “the nature of things” . . . . .	284

<i>CONTENTS</i>	xi
A.4.3 Abelard's arguments for the 3-formality of syllogisms . . . . .	285
A.5 Formal and material consequence . . . . .	292
<b>BIBLIOGRAPHY</b>	<b>297</b>

# List of Figures

3.1	Three notions of logical formality. . . . .	66
4.1	Kant's taxonomy of logics. . . . .	83
4.2	The argument for Kant's Thesis. . . . .	125
6.1	Components of the semantic enterprise. . . . .	188
6.2	The logical lattice for a four-valued logic. . . . .	212
6.3	The logical lattice for a nine-valued logic. . . . .	213
6.4	Four values collapsed into three. . . . .	230

# Chapter 1

## INTRODUCTION

### 1.1 What is logic?

From the instructor's perspective, the most difficult lecture of an introductory logic course is the first. The students quite naturally expect to be told what the discipline they will be studying is *about*. But the most obvious answer to this question—that logic studies relations of implication and compatibility between claims—does not bear up under reflection.

To be sure, logic *does* concern itself with relations of implication and compatibility between claims. The problem is that it is not the *only* science that does so. A logician can tell you that

(1) Julie is a girl who loves every boy.

implies

(2) Every boy is loved by some girl.

But to learn whether

(3) This substance turns litmus paper red.

implies

(4) This substance is an acid,

one has to ask a chemist, not a logician. It will be generally agreed, I presume, that the chemist investigating this question is not doing logic.

Shall we say, then, that logic concerns itself with *some* relations of implication and consistency, but not others? If so, then we cannot characterize logic generally as the study of implication and consistency relations. We will need to distinguish the implication and consistency relations with which *logic* is concerned from those that fall under the purview of the other sciences.

### 1.1.1 Method, not subject matter?

Before we go too far down this garden path, it is worth noting an alternative. It may be that logic is distinguished from other disciplines not so much by its subject matter (implication, consistency) as by the *methods* it employs, so that it is not a concern with implication relations that makes one a logician, but the techniques one uses to study them. On this view, nothing about the inference from (3) to (4) *itself* makes it an inappropriate target for the logician's inquiry. It's just that the logician's characteristic *methods*—e.g., the use of syntactic systems to formalize proof and model-theoretic methods to study consequence—aren't much use in determining whether (3) implies (4).<sup>1</sup>

Before we get too far in the search for distinctively *logical* relations of implication and consistency, then, we ought to ask whether we are not mistaking a distinctive *method* for a distinctive subject matter. In his introduction to *Model-Theoretic Logics* (Barwise and Feferman 1985), Jon Barwise suggests that those who draw a line between “logical concepts” (i.e., the constants of first-order logic) and other mathematical concepts are

... [confusing] the subject matter of logic with one of its tools. FOL is just an artificial language constructed to help investigate logic, much as the telescope is a tool constructed to help study heavenly bodies. From the perspective of the

---

<sup>1</sup>Granted, the characteristic methods of logic can be applied to subject matters from the special sciences. But it would not be unreasonable to say, for example, that a physicist who uses techniques from proof theory to test the consistency of a proposed axiomatization of quantum theory is doing logic—just as she is doing mathematics when she solves a system of equations in order to generate the consequences of a physical hypothesis. A proper delineation of logic should not imply that only logicians can do it.

mathematician in the street, the FO thesis is like the claim that astronomy is the study of the telescope. (6)

Perhaps there is nothing special about the kinds of implication studied by logic beyond their amenability to systematization using the evolving methods of logic.

But if logic is characterized by its *methods*, not its subject matter, then we owe a characterization of logical *methods*. If *any* methods for studying implication count as logical, then we are back to the original inadequate account of logic as the study of implication. But we should not restrict logical methods to the proof-theoretic and model-theoretic methods in use today, which were after all invented to solve problems antecedently recognizable as logical. On the other hand, if we characterize logical methods broadly enough to include Aristotle's *Prior Analytics*, it will be difficult to exclude Euclid's *Elements*.

For now, I simply want to flag the possibility of this alternative approach and return to the original path. How might we distinguish the kind of implication relations studied by logic (e.g., the one that holds between (1) and (2)) from the others (e.g., the one that holds between (3) and (4))?

### 1.1.2 Enthymemes

Some will object that *no* implication relation holds between (3) and (4), on the grounds that the inference from (3) to (4) is incomplete or *enthymematic*. The role of the chemist, they will say, is simply to determine the truth of the auxiliary premise that would be required to turn the inference into a complete one, namely:

(5) Any substance that turns litmus paper red is an acid.

On this view, there is no problem with the original demarcation of logic as the study of implication and consistency. While other disciplines attend to the truth or falsity of claims, questions about which claims *imply* which others are the proprietary domain of the logician.

But this response presupposes the theoretical work it seeks to avoid. Surely there is a pretheoretical sense of implication—of what counts as a *reason* for what—in which (3)

implies (4). (Consider how ridiculous it would be to bring a logician into a court of law to dispute a chemist's testimony to this effect.) In refusing to acknowledge that (4) follows from (3), the objector is deploying a *technical* sense of implication. And the task of distinguishing this technical sense of implication from the pretheoretical one is essentially the same as the task of saying *which* relations of implication (in the pretheoretical sense) fall under the purview of logic. Thus the enthymematic ploy simply *re-describes* the work that must be done in demarcating logic; it does not avoid this work.

### 1.1.3 Analyticity, a priority, necessity

It might be thought that *logical* implication can be characterized as implication that is knowable *a priori*, necessary, and/or analytic. (Alternatively, we might say that logic is the study of implication in the strict sense, and that implication in the strict sense must be knowable *a priori*, necessary, and/or analytic.) All three criteria seem to weed out the inference from (3) to (4). We cannot know *a priori* that (3) implies (4): we must do *experiments* in order to see that there is a connection between turning litmus red and acidity. Nor is the inference from (3) to (4) *necessary*, in any deeper sense than “demanded by physical law”: if the laws of nature were different, acids might not turn litmus red. Finally, our *understanding* of the meanings of “acid,” “litmus paper,” and “red” does not give us a sufficient basis for accepting that (3) implies (4).<sup>2</sup>

But there are three problems with characterizing logic as the study of *a priori*, necessary, and/or analytic implication. First, this approach presupposes that we have a reasonably clear understanding of *a priori* knowledge, (broadly logical) necessity, and/or analyticity—notions many philosophers regard as dubious or at least unclear.<sup>3</sup> Perhaps the very idea

<sup>2</sup>“Litmus” is the name for a particular substance derived from lichens; it is not *defined* as “any substance that turns red in contact with an acid.”

<sup>3</sup>Logical necessity is often thought to be less problematic than analyticity or *a priori* knowledge, because we can give precise proof-theoretic and model-theoretic definitions of logical truth and consequence. But these definitions presuppose that the demarcation of logic (e.g., the distinction of constants or inference rules into logical and non-logical) has already been accomplished, so they cannot provide a basis for it. Some philosophers (e.g., McFetridge 1990, Hale 1999) have tried to characterize logical necessity in a way that does not presuppose a prior demarcation of logic, but I am not convinced that their attempts are successful.

that there is a non-arbitrary dividing line between logic and other disciplines stands or falls with these notions. But before drawing that conclusion, it seems worth exploring other alternatives.

Second, such demarcations threaten to make the scope of logic much broader than it has traditionally been taken to be. For example, the inference from

(6) The ball is red.

to

(7) The ball is colored.

is surely analytic, *a priori*, and necessary, if anything is; but it is usually not held to be *logically* valid.<sup>4</sup> Again, if the natural numbers exist necessarily, then

(8) There are infinitely many objects.

is a necessary truth; but it is often held that logic alone cannot justify any existence claims. I will not multiply examples. This objection is not merely an appeal to tradition. There are important philosophical positions—notably Kantianism—that cannot even be *articulated* unless there is conceptual space between the logical and the necessary and *a priori*.

Third, while it may be true that logic is analytic, knowable *a priori*, and necessary, it would be nice if we could *explain why* logic has these features. By building the notions of analyticity, *a prioricity*, or necessity into the very *concept* of logic, we preclude meaningful discussion of whether and why logic is analytic, knowable *a priori*, or necessary. We also preclude using the notion of logic in an account of analyticity or necessity, on pain of circularity. Yet historically it has usually been the notion of *logic* that elucidates analyticity

---

<sup>4</sup>Etchemendy 1990 bites the bullet and accepts inferences of this sort as logically valid. Once we see the conceptual problems with the Tarskian account of logical consequence, he thinks, we lose the motivation for restricting logic to the study of implications that depend on the meanings of a small set of logical constants. “It is a mistake to think that . . . the consequence relation that arises from the meanings of predicate or function terms is any less significant than the logic of connectives and quantifiers. Once again, it is only the conflation of logical consequence with model-theoretic consequence that inclines us to think otherwise” (158). Chihara 1998 points out that in Etchemendy and Barwise’s computer program *Tarski’s World*, the sentence “for all x and y, if x is to the left of y, then y is to the right of x” is given as an example of a “logically valid” sentence (163).

and necessity, not vice versa. For example, Kant appeals to logic to clarify the difference between analytic and synthetic judgments (KrV:A151/B190), Frege defines analytic truths as those whose justification requires only logic and definitions (FA:§3), and Leibniz invokes the notion of logical compatibility in explaining the notion of a possible world.<sup>5</sup>

For all of these reasons, we ought to seek a demarcation of logic that does not appeal to the notions of analyticity, *a prioricity*, or necessity.

#### 1.1.4 Formal vs. material

How else might we draw a line between the implication from (1) to (2) and the implication from (3) to (4)? The tradition suggests another ground for the distinction: the first is *formal*, the second *material*. William Kneale 1956 appeals to this distinction in criticizing his contemporaries' penchant for speaking of “the logic of color words” or “the logic of psychological words”:

With them “logic” is no longer the name of a science concerned with the principles of inference common to all studies, but rather a name for any collection of rules in accordance with which we may argue in some context. One philosopher of this persuasion has even said that every kind of statement has its own logic. In such a welter of metaphor and epigram it is difficult to know what we are expected to take seriously; but it seems clear that this way of talking involves abandonment of the notion that logic is concerned with form as opposed to subject matter. (238)

But what does it *mean* to say that logic is concerned with form rather than matter or content, or that it is “topic-neutral,” “independent of subject matter,” or distinctively “formal”? That is the topic of this investigation. Although the topic is of contemporary interest, my approach is largely historical. In what follows, I examine the tradition of characterizing logic as distinctively *formal*—a tradition I call *logical hylomorphism*—and distinguish (a) different *notions* of logical formality that are in play and (b) different *justifications* for the claim that logic must be distinctively formal in one or more of these senses. The aim is to

---

<sup>5</sup>An exception to this usual order of explanation is Carnap, who uses the notion of analyticity to explicate logicity and necessity (1934, 1937, 1942, 1947).

bring some clarity to present debates about the demarcation of logic by making it possible to explain and systematize our intuitions about logicity.

But why do we need a demarcation of logic at all? And why focus on *formality*? Finally, why is it appropriate for the investigation to be historical? In the remainder of this chapter, I motivate the project by answering these questions.

## 1.2 Why do we need a demarcation?

Why is it necessary to draw a principled distinction between logic and non-logic at all? If all that were at stake were the labels on office doors or the titles of journals, then we could rest content with an appeal to “family resemblance” (Wittgenstein 1953:§66-7). We could afford to be inclusive about what we counted as logic, because little would depend on our classification. As Steven Wagner 1987 observes, “Trying to characterize logic is pointless unless one has a viewpoint from which it matters what logic is” (7). It is therefore worth examining some philosophical projects for which the demarcation of logic (in some suitably narrow sense) really *matters*.

### 1.2.1 Logicism

The theses of logicism—that mathematical concepts are definable in terms of logical concepts, that mathematical truths are reducible to logical truths, that mathematical modes of inference are reducible to logical modes of inference, and that mathematical knowledge is really logical knowledge—are philosophically interesting only to the extent that there is something special about logical concepts, truths, modes of inference, and knowledge. In the absence of some view about the special character of logic, logicism is no more philosophically significant than any other definability thesis in mathematics. Curry 1951 puts the point well:

[Logicism] is said to be characterized by the fact that it reduces mathematics to logic. This appears to be a thesis in regard to the definition of mathematical truth. On closer examination, however, it is evident that we do not have

a theory of mathematical truth parallel to those already considered [i.e., intuitionism, formalism, Platonism]. For, so long as ‘logic’ is undefined, to say that mathematics is logic is merely to replace one undefined term by another. When we go back of the word ‘logic’ to its meaning we find that the logicians have very varied conceptions of logic and so of mathematics. (65)

Logicism would be trivial if it could be defended by saying: “I take as my underlying logic the whole of analysis, together with its characteristic modes of inference. Now every theorem of analysis can be derived without appeal to any extra-logical axioms. QED!”<sup>6</sup>

Accordingly, a popular way to attack logicism is to argue that the “logic” to which mathematics has been reduced is really just more mathematics. Poincaré 1908 charges that Russell’s “logical” principles are really intuitive, synthetic judgments in disguise:

We regard them as intuitive when we meet them more or less explicitly enunciated in mathematical treatises; have they changed character because the meaning of the word logic has been enlarged and we now find them in a book entitled *Treatise on Logic?* (1946:461)<sup>7</sup>

Quine, a more sympathetic critic, draws a similar conclusion:

Frege, Whitehead, and Russell made a point of reducing mathematics to logic; Frege claimed in 1884 to have proved in this way, contrary to Kant, that the truths of arithmetic are analytic. But the logic capable of encompassing this reduction was logic inclusive of set theory. (1986:66)

Applied to the full-fledged logicism that identifies numbers with classes, Quine’s criticism is just—and nearly universally accepted.<sup>8</sup>

It is not so clearly decisive, however, against the more limited or retrenched forms of logicism that are sometimes defended today. Boolos 1985 defends a thesis he calls “sublogicism”: the view that *many* significant mathematical concepts and truths can be reduced to

<sup>6</sup>Jané 1993 notes that there is a broad sense of “logic” “. . . in which *every* language has its logic, determined by its consequence relation, regardless of how much content is carried by it. . . . Indeed, any mathematical theory can be embedded in the logic of some language just by treating the terms peculiar to the theory as logical particles of the language . . .” (67). This cannot be the sense of “logic” at issue in discussions of logicism.

<sup>7</sup>On Poincaré’s criticisms of logicism, see Parsons 1965 and Goldfarb 1988.

<sup>8</sup>Frege’s class existence axiom, a plausible candidate for a logical law, was shown to be inconsistent, and all known consistent theories of classes require existence assumptions that seem to lack the “special character” of logic. For Frege, this spelled the death of logicism. Although Russell persisted in maintaining the logicist theses, he acknowledged that his axioms of infinity and reducibility are not logical and must be taken as hypotheses (1920:141, 193).

logical ones. There is no denying that sublogicism is philosophically interesting: if true, it suffices to demolish the Kantian view that all significant mathematics depends essentially on a non-logical basis, our *a priori* forms of intuition (156).<sup>9</sup> Against sublogicism, it can be urged that the second-order logic it employs is not really logic, but “set theory in sheep’s clothing” (Quine 1986:66), and that *calling* it logical does not give it any special epistemic status. This is not the place to enter into the debate over second-order logic (see Quine 1986, Boolos 1975 and 1984, Wagner 1987, Resnik 1988, Shapiro 1991, Jané 1993). My point is this: if it matters whether sublogicism is true, then it matters whether second-order logic is really logic, and thus it matters how logic is demarcated.

Another kind of retrenched logicism springs from Wright 1983. The technical discovery on which Wright’s “number-theoretic logicism” rests is that the Peano Axioms for arithmetic can be deduced from the laws of higher-order logic with identity and the single non-logical axiom

( $\mathbf{N}^=$ ) The number of Fs = the number of Gs iff the Fs and the Gs can be put in a one-one correspondence.

The right side of this biconditional can be expressed in pure second-order logic; the only non-logical vocabulary in  $\mathbf{N}^=$  is the functor “the number of.” For this reason, Wright holds that  $\mathbf{N}^=$ , though not an explicit definition of number in purely logical terms, can be regarded as an *explanation* of the concept of cardinal number in purely logical terms (153). The upshot, he claims, is that arithmetic is analytic, since every arithmetical truth is a logical consequence of the explanation of a concept. Wright has come under heavy fire for this claim.<sup>10</sup> What is relevant for our present purposes is that in order to defend this position, Wright needs to argue that the second-order quantifiers used in  $\mathbf{N}^=$  and in the derivations of the Peano Axioms from  $\mathbf{N}^=$  are genuinely *logical* quantifiers (132-5)—

<sup>9</sup>In support of sublogicism, Boolos adduces Frege’s derivations of the propositions that the ancestral is transitive and that the ancestral of a function is connected (157). Boolos is right, I think, that the content of these propositions “... can be seen as a generalization of that of familiar and fundamental mathematical principles, for the grasp of whose truth some sort of ‘intuition’ was often supposed in Frege’s time to be required” (158).

<sup>10</sup>See Field 1984 and the essays in Demopoulos 1995 and Boolos 1998.

quantifiers that are essential components, as he puts it, of *formally* valid inferences. He makes some moves toward discharging this obligation, but not without acknowledging its difficulty and importance:

Quite how, or indeed whether, the idea of a formally valid inference can be made good is a question of the greatest difficulty; we can decently evade it here only because that there are such inferences is presupposed by the very existence of logic as a special science, and because the logicist's thesis cannot be so much as formulated unless there is such a special science. (133)

Just my point.

Hartry Field 1989 defends the claim that mathematical knowledge is ultimately logical knowledge: knowledge of which mathematical claims follow from which, and of which bodies of mathematical claims are consistent. Because he denies that mathematical truths are logical truths—indeed, he denies that mathematical claims are true at all—he eschews the label “logicism,” calling his view “deflationism” (81). But the view is essentially the same as the “conditional logicism” (Coffa 1981) that Russell advocates in the first part of Russell 1903. As in the cases of retrenched logicism, the view is interesting only to the extent that the logic used has a special character—in this case, a special epistemic character. Deflationism is pointless unless our knowledge of logic is less problematic than our knowledge of mathematics (taken at face value).<sup>11</sup> Field's justification for this claim rests on a substantive view about logic which he takes from Kant: the view that logic can make no assertions of existence (80).

In sum, there are two ways in which the demarcation of logic matters for these projects (full logicism, sublogicism, number-theoretic logicism, and deflationism). First, the *technical* success of each project depends on whether the machinery it requires (e.g., second-order quantification) counts as logical. Second, the philosophical significance of a (technically successful) version of logicism depends largely on the philosophical significance of the concept of logic that is in play. If logic turns out to have no interesting epistemic, modal, or semantic properties, then logicism cannot have the significance it has been thought to have.

---

<sup>11</sup>For a criticism of Field's claim to have made an epistemic advance, see Shapiro 1993.

### 1.2.2 Structuralism

If you asked mathematicians in 1800 what the subject matters of geometry, arithmetic, and algebra were, they would have answered with little hesitation: *extension* and (determinate and indeterminate) *quantity*. By 1900, however, many mathematicians were taking geometry, arithmetic, and algebra to be about *abstract structures* implicitly defined by formal axiom systems.<sup>12</sup> This change was driven largely by changes in the mathematics, not philosophical reflection. Negative and imaginary numbers, which proved indispensable in algebra, were difficult to conceive as *quantities*. In order to accommodate them, algebraists began conceiving of their subject as the study of abstract operations defined by explicit laws. Boole's friend and mentor Duncan Gregory articulated the new view as follows:

The light in which I would consider symbolical algebra is that it is the science which treats of the combination of operations defined not by their nature, that is, by what they are or what they do but by the laws of combination to which they are subject. (1838:208, quoted in Nagel 1979:183)

So conceived, the theorems of algebra apply not just to addition and multiplication of numbers, but to “any operations in any science [that are] subject to the same laws of combinations” (ibid.).

A similar perspective was eventually achieved in geometry, though it was slower and harder won. Projective geometers invoked “imaginary elements” that could not be visualized in terms of extension. The discovery of the principle of duality in projective geometry—that for each theorem in projective plane geometry and its proof, there is “dual” theorem (and proof) in which “line” and “point” are interchanged—suggested that geometry was more general in its subject matter than had previously been allowed. Pasch took the principle of duality to show that proper geometric deductions must depend only on the “*relations* specified in the propositions and definitions employed,” and not on the (spatial) *meanings* of the geometrical concepts (1926:91, quoted in Nagel 1979:237). Only such deductions could

<sup>12</sup>For a masterful recounting of these changes, see Nagel 1979 (chapters 8 and 9), from which much of the following is taken. Shapiro 1997 (chapter 5) connects this history with contemporary structuralism in the philosophy of mathematics.

establish the theorem in its full generality. The fruit of this development was Hilbert's *Grundlagen der Geometrie* (1899), a careful and rigorous axiomatization of geometry that studiously avoids appeal to spatial intuition (except for purposes of motivation). Hilbert boasted that from the point of view of pure geometry, it is irrelevant whether the abstract structures defined by the axioms are instantiated by points, straight lines, and planes, or by tables, chairs, and beer mugs (Shapiro 1997:157): geometry concerns only the stipulated relations between these elements.

What really distinguishes this new view of pure mathematics from the more traditional view is not so much the emphasis on *relations* between elements, as opposed to the elements themselves—Kant, too, took mathematics to concern the *forms* of objects—as the demand that these relations be capable of being made explicit without any appeal to spatial or temporal intuition. Pure mathematics, as conceived by contemporary “structuralists” in the philosophy of mathematics, concerns only *pure* or “freestanding” structures—structures whose relations can be defined in purely logical terms.<sup>13</sup> The *natural number* structure, for instance, can be defined by the following axioms:

- (P1)  $\mathbf{N0}$
- (P2)  $\forall x(\mathbf{N}x \supset \mathbf{N}s(x))$
- (P3)  $\forall x\forall y((\mathbf{N}x \ \& \ \mathbf{N}y \ \& \ s(x)=s(y)) \supset x=y)$
- (P4)  $\sim \exists x (\mathbf{N}x \ \& \ s(x)=\mathbf{0})$
- (P5)  $\forall F( (\mathbf{F0} \ \& \ \forall x((\mathbf{N}x \ \& \ \mathbf{F}x) \supset \mathbf{F}s(x)) ) \supset \forall x(\mathbf{N}x \supset \mathbf{F}x) )$

The only non-logical terms in these axioms are  $\mathbf{N}$ ,  $\mathbf{0}$ , and  $s()$ , which are given their contents by the axioms alone.

Not every structure is freestanding in this way. A musical composition—say, Bach's D-minor Chaconne for solo violin—can be thought of as a structure, inasmuch as it can be instantiated by performances on different instruments and with different phrasings and tempi. Perhaps it can even be instantiated by a performance in a different key. But can

<sup>13</sup>For a survey of several varieties of structuralism in the philosophy of mathematics, see Shapiro 1997. For the concept of *freestanding* structures, see 100-106. For a mathematician's view on structure in mathematics, see MacLane 1996.

it be instantiated by a sequence of lines of different lengths? Or by a set of nested circles of various colors? Or by a single complex aroma? Granted, one can find a *more general* structure of which all of these are instances, but this more general structure is not the D-minor Chaconne structure. The D-minor Chaconne structure is *essentially* the structure of a musical composition: it cannot be defined without using relations of relative musical pitch and temporal duration. Hence it is not a *pure* or (in Shapiro’s terminology) *freestanding* structure, and it lies outside of the realm of pure mathematics.

On the structuralist account, then, pure mathematics is distinguished from other disciplines by its exclusive concern with freestanding structures. But since freestanding structures are structures defined solely in terms of *logical* relations, this way of demarcating pure mathematics presupposes a demarcation of logic.<sup>14</sup> For instance, if tense logic counts as logic (for the purpose of defining “freestanding”), then the structure defined by the following axioms will be an object of pure mathematics:<sup>15</sup>

- (T1) **Pa**  
 (T2)  $\forall x(\mathbf{P}x \supset \text{Was}:\exists y(\mathbf{P}y \ \& \ \mathbf{R}xy \ \& \ \sim y=x))$   
 (T3)  $\forall x\forall y\forall z((\mathbf{R}xy \ \& \ \mathbf{R}xz) \supset y=z)$   
 (T4)  $\sim\text{Will}:\exists x(\mathbf{P}x \ \& \ \mathbf{R}xa),$

where the only non-logical terms are **P**, **R**, and **a**. The problem is that the existence of this structure—not just empirical instantiations of the structure, but (for realists about structures) the structure itself—involves substantial assumptions about the nature of time (e.g., the existence of a past history with an infinite number of distinct moments), assumptions that (on most views) lie outside the scope of pure mathematics.

What this example shows is that the philosophical significance of the distinction between pure and applied mathematics depends on the way in which logical relations are singled out. A strong historical motivation for the move towards structuralism was the desire to purify the objects of mathematics of any tincture of intuition or sensibility. If the structuralist

<sup>14</sup>Cf. Simons 1998:499: “It is important to be able to state what a logical property is without invoking the notion of a pure structure, otherwise circularity will result.”

<sup>15</sup>I use “Was:” and “Will:” for the standard past- and future-tense operators.

account of pure mathematics is to satisfy this desire, then the demarcation of logic on which it is based must show *logic* to be independent of sensible intuition. At the same time, it must yield a logic strong enough to characterize the important structures studied by pure mathematics, which means going beyond standard first-order logic. In this way, the structuralist account of pure mathematics both depends on and gives point and purpose to the demarcation of logic.

### 1.2.3 Explicitation

In axiomatizing an informal theory, we seek to make explicit the assumptions on which it depends and the concepts it invokes. We do this for practical purposes as well as philosophical ones. Separating out a set of axioms from which the whole theory unfolds logically allows us to focus our assessment of the theory and its justificatory grounds on the axioms alone. Axiomatization also makes it possible to study the relations between alternative theories: relative consistency, relative strength, and so on. The machinery of mathematical logic has been developed largely to further projects like these, not the explicitly philosophical projects with which we have recently been occupied.

The theoretician's use of logic as a framework for explicitation imposes its own constraints on the delineation of logic. It is simplest to illustrate the point with an example, due to Jané 1993. Jané defines a quantifier  $Q$  as follows:

$Qx\alpha(x)$  is true in a structure iff  $\langle A, +, \times \rangle$  is isomorphic to the field of real numbers (where  $A$  is the set of elements of the domain of the structure satisfying  $\alpha$ , and  $+$  and  $\times$  are the interpretations of  $+$  and  $\times$ , respectively). (73)

Adding  $Q$  to first-order logic as a logical constant, we obtain a new logic, "*R*-logic." *R*-logic might serve as a useful background logic for the axiomatization of a theory in which the reals could be taken for granted—e.g., the theory of real vector spaces. Such an axiomatization would make explicit the salient assumptions of the theory of real vector spaces, but assumptions about scalars would remain implicit in the underlying logic. Indeed, a single sentence of pure *R*-logic,  $Qx(x=x)$ , provides "... a complete and categorical axiomatization

of the real field” (73). Clearly, although we can *characterize* the structure of the real field using  $R$ -logic—in the sense that we can pick out the relevant class of models— $R$ -logic is not the proper framework for its *explicitation*.

For similar reasons, Jané argues, second-order logic is not the proper framework for the study of set theory. There is simply too much set theory implicit in the second-order consequence relation. For example, the explicit formulation of the Axiom of Choice is surely one of the triumphs of early axiomatic set theory: we learned something important about our conception of *set* when we made this assumption explicit. But Choice need not be stated at all in a second-order axiomatization of set theory (83). It has been absorbed, so to speak, into the underlying logic.

If set theory is to serve as a foundation for all mathematics, Jané concludes, it must be investigated in an underlying logic free of substantive mathematical content (67, 83). Jané takes first-order logic to be such a logic, on the grounds that it has a complete proof procedure (67)—but that is another argument. What concerns us now is simply the way in which the *role* of logic as a framework for the explicitation of the content of theories makes it *matter* how logic is demarcated.<sup>16</sup>

### 1.3 Why formality?

Granting, then, that we have reason to seek a demarcation of logic, how should we approach the task? There are many possibilities here. This dissertation examines the idea that logic is distinguished from other disciplines by its *formality*. (What that comes to will be considered in chapter 3.) It is worth locating and motivating this strategy in relation to some of the other possible approaches to the problem.

---

<sup>16</sup>Similar issues arise in the context of *philosophical* theorizing. For example, Quine argues that the use of an underlying logic with branching quantifiers disguises ontological commitments (to the existence of functions) that first-order logic would force to be made explicit (1986:89-91; 1969:chapter 4; cf. Wagner 1987:17-19). Similarly, Shapiro 1998a accuses semantic deflationists who use strong underlying logics of hiding the “robustness” of truth in the strong consequence relation.

### 1.3.1 Pragmatic vs. principled demarcations

First, it is useful to distinguish two general strategies for demarcating logic. The first is to identify some favored property (formality, topic-neutrality, conservativeness of introduction and elimination rules, permutation invariance, etc.), or perhaps some combination of properties, as a necessary and sufficient condition for logicity.<sup>17</sup> Call approaches of this kind *principled demarcations*. The second strategy is to start with a particular *job*—for example, to serve as a “framework for the deductive systematization of scientific theories” (Warmbröd 1999:516), or a framework for the characterization of structures—then find something that is capable of doing that job, and identify it as logic. Call approaches of this kind *pragmatic demarcations*.

The literature on demarcating logic contains examples of both general approaches,<sup>18</sup> and it is important to recognize three differences between them. First, pragmatic approaches are guided by what Warmbröd calls a “requirement of minimalism”:

...logical theory should be as simple, as modest in its assumptions, and as flexible as possible given the goal of providing a conceptual apparatus adequate for the project of systematization. In practice, the minimalist constraint dictates that the set of terms recognized as logical constants should be as small as possible. (Warmbröd 1999:521)

Or, in Harman’s pithy formulation: “Count as logic only as much as you have to” (1972:79). Warmbröd uses this constraint to argue that the theory of identity is not part of logic, on the grounds that “[w]e can systematize the same sets of sentences by recognizing only the truth-functional connectives and first-order quantifiers as constants, treating ‘=’ as an ordinary predicate, and adopting appropriate axioms for identity” (521).<sup>19</sup> On similar grounds, both Harman and Warmbröd urge that modal operators should not be considered part of logic.<sup>20</sup>

<sup>17</sup>A condition for *what* to be logical? Concepts, rules, implications, systems, constants? Here I leave this question open. Different demarcation projects answer it differently.

<sup>18</sup>*Principled*: Kneale 1956, Tarski 1966, Dummett 1973:21-3, Peacocke 1976, Hacking 1979, McCarthy 1981, Sher 1991 and 1996, Došen 1994 (for a useful overview, see Sainsbury 1991:312-25). *Pragmatic*: Harman 1972, Tharp 1975, Quine 1986:ch. 5, Wagner 1987, Shapiro 1991, Warmbröd 1999.

<sup>19</sup>For more discussion of this proposal, see section 2.3.2, below.

<sup>20</sup>Warmbröd’s approach is to paraphrase modal claims into first-order claims about relations

Their point is not that identity or modal operators lack some characteristic that the first-order quantifiers and truth-functional operators possess, but merely that, since we *can* get by without taking these notions to be part of our logic, we should. Warmbröd and Tharp even explore the possibility of taking truth-functional logic to be the whole of logic and viewing quantification theory as a non-logical theory (Warmbröd 1999:525, Tharp 1975:18), though both reject this idea on pragmatic grounds.

While pragmatic demarcations seek to *minimize* what counts as logic, principled demarcations are *inclusive*. They count as logical *any* notion (rule, system, etc.) that has the favored property. It is simply irrelevant whether a notion (rule, system) is required for a particular purpose: its logicity rests on features that it has independently of any use to which we might put it.

A second difference is that pragmatic approaches tend to be *holistic*, in a way that principled approaches usually are not. Since it is the whole logical *system* that does the job, systemic properties—decidability, completeness, compactness, the Löwenheim-Skolem property—play a greater role in pragmatic demarcations than local features of particular notions or inference rules (see Quine 1986:ch. 5, Tharp 1975, Wagner 1987). Although nothing about the idea of a principled demarcation excludes appeal to systematic properties, principled demarcations tend to appeal to local properties of particular notions or inference rules (e.g., topic-neutrality, necessity, introducibility through conservative inferential definitions), rather than systemic properties.

The third (and most important) difference between pragmatic and principled demarcations has to do with the *consequences* one can draw from their verdicts about logicity. A demarcation tells us that *this* implication is a logical consequence, *that* one non-logical: but what *follows* from this classification? Not much, if the demarcation is pragmatic; potentially quite a bit, if the demarcation is principled.

---

between possible worlds, while Harman’s is to introduce a non-logical predicate “is necessary” and a logical operator that forms names of propositions. As Steven Kuhn 1981 shows, it is also possible to take the sentential operators “necessarily” and “possibly” as non-logical operators in an intensional language.

To see why, consider the significance of a demarcation of logic for the evaluation of logicism. What would be the upshot of showing that some significant part of mathematics is reducible to logic, where logic is given a *pragmatic* demarcation? We would have shown that the minimal conceptual apparatus that is sufficient for some purpose (say, deductive systematization of theories) already allows us to do a considerable amount of mathematics. Such a result would not be uninteresting, but it would not shed much light on traditional questions in the philosophy of mathematics (what is the source of the objectivity of mathematics?; does mathematics have special objects?; how do we come by our mathematical knowledge?). If, on the other hand, we had a *principled* demarcation of logic, then a logicist thesis would have clear philosophical significance: whatever privileged feature we use to pick out logic could be immediately transferred to the reduced body of mathematics. For example, if logical implication were demarcated in part by its *necessity* (in some sense), then the reduced body of mathematics would have been shown to be necessary (in the same sense).

The point I am making does not depend on whether the logicist thesis in question is shown to be true or to be false. Given a principled demarcation, a demonstration of the *falsity* of the logicist thesis would show that some body of mathematics *fails to have* the feature that distinguishes logic. But given a pragmatic demarcation, all the falsity of a logicist thesis amounts to is that not all of the conceptual resources required to do mathematics are required for the fundamental task of logic. There is no implication that these conceptual resources lack some feature that logical resources possess.

Indeed, on the pragmatic approach, what counts as logic depends on the current state of scientific and mathematical theory. If the advance of science results in an increase in the resources needed for deductive systematization (or whatever is the favored task of logic), then these resources automatically count as logical (Warmbröd 1999:533). On a principled approach, by contrast, whether particular resources are logical depends only on whether they have the favored property. If they do not, and if it turns out that they are needed for the deductive systematization of theories, then the proper conclusion to draw is that logic

alone is not adequate for this task. (Kant, as we will see, was quite happy to draw this conclusion.)

It seems to me that the philosophical projects that give point and purpose to demarcating logic in the first place (see section 1.2, above) require a *principled* rather than a pragmatic demarcation. The philosophical interest of projects like logicism depends on logicity having some philosophically interesting essence—hence on logic being picked out by a distinguishing property or properties, not by its suitability for some task. It is not clear what philosophical *purpose* is served by pragmatic demarcations. Hence the focus of this investigation is on principled demarcations.

Indeed, the focus is even narrower—on demarcations of logic by its “formality.” What motivates this choice? Many articles on the bounds of logic make no use of the hylomorphic terminology (a prudent practice, given the pitfalls I will be mapping in what follows). Why should “formality” be the focus of our inquiry? It will be helpful in answering this question to consider how *else* one might proceed in investigating the demarcation of logic.

### 1.3.2 Alternatives

Many demarcation projects invoke a *technical* property: invariance under all permutations of a domain of objects, for instance, or definability by introduction and elimination rules having a certain form and conservatively extending a given inferential base. But approaches of this kind must *justify* their choice of technical property: it is not just obvious that conservative introducibility or permutation invariance should be relevant to defining logic. And it is not a sufficient justification to show that the technical property picks out a set of notions or rules that we antecedently regarded as logical. If *that* is all we have, then our demarcation of logic will be merely a codification of our intuitions about particular cases; and if the enterprise of demarcation is to have a point, it must provide more than just that. Besides, there is little agreement about logicity in particular cases. Hence, in order to justify a *technical* demarcation of logic, one must appeal to some *non-technical* characterization of logicity—for instance, the characterization of logic as *formal* (cf. Sher

1991, Došen 1994).

Granted, one need not appeal to the *formality* of logic. But what else is there? One might invoke the modal and/or epistemic features of logical truth and consequence (necessity, a prioricity). But we have already seen reasons to shun this approach (section 1.1.3, above). The other non-technical property of logic that is often invoked in demarcation projects is its distinctive *generality* or *topic-neutrality*. Logic is not supposed to be about anything in particular; it is distinguished by its lack of any special subject matter. But (as I will show in chapter 3) generality and topic-neutrality turn out to be unclear in precisely the same ways as formality. Thinking through the senses in which logic might be said to attend only to “form,” abstracting from subject matter or content, and thinking through the senses in which logic might be said to be “general” or “topic-neutral” are really the same enterprise. All of these terms have the same range of meanings.

It appears, then, that if one wants to give a principled demarcation of logic that does not rely on epistemic and modal characterizations of the discipline, one must make sense of what it means to say that logic is *formal*.

### 1.3.3 Historical importance of formality

A somewhat different argument for focussing on formality starts from its *historical* centrality in conceptions of logic. Consider just a few representative passages:

... the universal and necessary rules of thought in general can concern merely its *form* and not in any way its *matter*. Accordingly, the science that contains these universal and necessary rules is merely a science of the form of our cognition through the understanding, or of thought. (Kant JL:12)

The examples which we have hitherto employed lead naturally to a first principle part, which, under the name of pure or formal logic, is devoted to thought in general and those universal forms and principles of thought which hold good everywhere, both in judging of reality and in weighing possibility, irrespective of any difference in the objects. (Lotze 1843:8)

Logic inquires into the form of thought, as separable from and independent of the matter thought of. (De Morgan 1858:75)

... if there is one point on which logicians are agreed, it is that logic is formal, and pays no regard to anything not formally expressed. (Jevons 1864:69)

What is of concern to logic is not the special content of any particular relation, but only the logical form. (Frege GL [1884]:§70)

Thus the absence of all mention of particular things or properties in logic or pure mathematics is a necessary result of the fact that this study is, as we say, ‘purely formal.’ (Russell 1920:198)

The progress achieved by axiomatics consists in its having neatly separated the logical-formal from its objective or intuitive content; according to axiomatics the logical-formal alone forms the subject-matter of mathematics, which is not concerned with the intuitive or other content associated with the logical-formal. (Einstein 1921:28)

... since we are concerned here with the concept of logical, i.e., *formal*, consequence, and thus with a relation which is to be uniquely determined by the form of the sentences between which it holds, this relation cannot be influenced in any way by empirical knowledge, and in particular by knowledge of the objects to which the sentence X or the sentences of the class K refer. The consequence relation cannot be affected by replacing the designations of the objects referred to in these sentences by the designations of any other objects. (Tarski 1936:414-5)

... the derivation must be made without any reliance on geometrical assertions other than those taken as primitive, i.e. it must be *formal* or independent of the special subject matter discussed in geometry. (Kneale and Kneale 1962:4)

By far the most serious objection against Aristotle’s solution lies, however, in the fact that it entails an infraction on the *formal character* of logic; for in many cases the answer to the question whether a certain term is ‘empty’ or not is dependent on empirical data. (Beth 1970:124-5)

The second class is much too heterogeneous. It includes ‘bachelor’-‘male’ inferences, as well as e.g. the logical inferences. But these ought to be distinguished, since the latter are clearly *formal* in a sense in which the former are not. (Evans 1985:405)

I stressed before that in the proposal I have made the logical consequence relation, in opposition to the more general analytic consequence relation, has indeed a ‘formal’ character. That formal character derives from the ‘contentless’ meaning of the logical expressions... (Sánchez-Miguel 1993, 125)

These passages exemplify a tradition, going back at least to Kant and encompassing philosophers and logicians with wildly different philosophical views, of demarcating logic by its formality. I call this tradition *logical hylomorphism*. Logical hylomorphism is a tradition or current, not a *thesis*, because its advocates understand the formality of logic in very different ways. (In this respect it resembles traditions like liberalism and eudaimonism.) But it would be wrong to think that the tradition is held together only by a word. It is

held together in part by its members' intentions to be characterizing logic in accord with the tradition, despite their philosophical differences.<sup>21</sup>

It is remarkable that this tradition has never been adequately studied, and it would not be hard to motivate an investigation of it on purely historical grounds. But my interests here are primarily philosophical, and here one might be sceptical: "So what if there's a tradition of demarcating logic by its formality? Why should we philosophers trouble ourselves with it? We want to know what logic *is*, not what people have *said* about it. The tradition has no authority over *us*! We can do as we please."

## 1.4 Why a historical approach?

But *can* we do as we please? We can certainly *stipulate* an arbitrary definition for "logic," just as we could stipulate that henceforth we will use the word "giraffe" to designate cows. But stipulating a meaning for "logic" is not the same as saying what logic is.

Still, it may seem mysterious why an investigation of the nature of logic should require detailed attention to our traditional conceptions of logic. After all, one does not investigate the nature of *gold* by attending to our predecessors' *conceptions* of gold. Our predecessors believed all kinds of false things about gold (e.g., that it could be obtained by alchemical transformation from many other substances). And they lacked any grasp of what we now regard as the *defining feature* of gold: the number of protons in the nucleus of gold atoms. Studying our predecessors' beliefs about the essence of gold is next to worthless if we want to know what distinguishes gold from other substances. If we really want to know what gold is, we need to study *gold*. If we want to know what logic is, then, shouldn't we study *logic*, rather than *conceptions of logic*? Isn't there every reason to believe that our predecessors' conceptions of logic are as unreliable as their conceptions of gold?

This apparently cogent line of argument slurs over an important disanalogy between

---

<sup>21</sup>For example, in the *Principles of Mathematics*, Russell claims to have refuted Kant's view that mathematical proof requires non-formal reasoning (1903:§4, §§433-4), but the issue has only been joined if Russell means the same thing as Kant did (or close enough) by "formal."

the concept of *gold* and the concept of *logic*. Because *gold* is a natural kind concept, our grasp of it (even our best experts' grasps) can be partial and confused. As long as we can identify some paradigm samples of gold, and as long as we conceive of gold as a single substance, all of whose instances behave similarly in the natural order, we have managed to attach our thought to *gold* (and not to, say, the disjunctive kind *gold-or-pyrite* or the phenomenal kind *golden-colored metal*).<sup>22</sup> Nature herself makes up for the slack in our collective understanding. That is why we can investigate the nature of gold without worrying too much about how our predecessors conceived of it. All we need to know is that we are experimenting on and theorizing about the very same stuff to which they applied the concept *gold*.

Investigating the nature of logic, I suggest, is not like this. *Logic* is not a natural kind concept. It does not play a role in laws of nature, and so the natural order of the world cannot take up the slack between our ways of grabbing onto the concept and the concept itself. Thus we must attend more closely to the ways in which our predecessors marked out the subject if we want to ensure that we are talking about the same thing. This is not to say that there is no room to criticize our predecessors' conceptions. We want to be able to say that our predecessors were wrong (at least in part) about both the scope and the essential characteristics of logic. But our model for such criticism cannot be our criticism of our predecessors' conceptions of *gold*. The concept of *logic*, I suggest, is more usefully compared with *legal* concepts such as *negligence*, *property*, or *contract*. The correct application of these concepts requires much more sensitivity to past usage and past theory than does the correct application of *gold*. That is why judges must attend studiously to precedent. And that is why an investigation of the nature and bounds of logic must attend to the tradition of demarcating logic.

A philosopher with platonistic leanings in the philosophy of mathematics might object that it is chauvinistic to restrict *natural* kinds to kinds studied by the empirical sciences. Surely we can distinguish "natural" and "unnatural" kinds in mathematics, too. It is widely

---

<sup>22</sup>The point is now commonplace in the literature, but the *locus classicus* is Putnam 1975.

held, for instance, that the convergent analyses of effective calculability given by Gödel, Kleene, Church, and Turing amount to a *discovery* of the nature of effective calculability, in much the same way as the atomic theory made possible the discovery of the nature of gold. Might not mathematical results also show us how to demarcate logic? Kneale and Kneale 1962 suggest that Gödel's incompleteness results show that "logic extends no further than [first-order] quantification theory":

When Frege wrote, the scope of logic had not been delimited precisely, and his [logician] thesis seemed plausible just because the reader could then make an easy transition in thought from quantification theory to the theory of sets and arithmetic. But Gödel has revealed a profound difference between quantification theory, which is complete, and the theory of sets, which is not. In the interests of clarity it therefore seems best to reserve the name "logic" for the former, and this is in fact what most mathematicians do when they are engaged upon their ordinary concerns. (724; cf. 741)

Perhaps, then, it is the mathematical *logician*, not the philosopher of logic, who is best placed to tell us what logic is—just as it is the *chemist*, not the philosopher of science, who is best placed to tell us what gold is.

I do not want to deny that technical results like Gödel's can reveal natural conceptual joints. But one should not overplay the analogy between chemistry and mathematical logic. The concept of *logic* plays as important a role in philosophy as it does in mathematics (perhaps a *more* important role). So although mathematical results are *relevant* to the demarcation of logic, they cannot bear the whole burden. Even the Kneales appeal to the tradition (specifically the tradition of taking the logical enterprise to be that of "classifying and articulating the principles of formally valid inference," 739) to support the demarcation they favor (741).

One might still resist my conclusion that an intelligent, principled demarcation of logic must be grounded in a thorough study of the *history* of conceptions of logic. But what are the alternatives? If we do not look to history, how do we know when we have gotten the *right* demarcation? I see only two possible replies, and neither, I will argue, is satisfactory.

The first reply is to reject the question. On this approach, there is no such thing as the

“right” demarcation. One can *construct* any concept of logicity one wishes—simply by stipulation. None is any better or worse than any other, except in relation to a particular purpose. That we call them concepts of “logicity” has only psychological significance.

To see what is wrong with this reply, try substituting “negligence” for “logicity.” Of course we can stipulate various meanings for “negligence,” but if that’s *all* we can do, then we lose something very important: continuity of subject matter with our intellectual (and legal) predecessors. We want to be talking about the same thing they were talking about, so we can profit from or correct their reasoning, and no amount of *stipulation* can ensure that we are doing so. Continuity of subject matter is especially important in connection with logicity, because the reason we *care* about the concept of logicity is the role it plays in debates in philosophy and the foundations of mathematics—ongoing debates, with histories. As Sellars 1967 puts the point,<sup>23</sup>

The history of philosophy is the *lingua franca* which makes communication between philosophers, at least of different points of view, possible. Philosophy without the history of philosophy, if not empty or blind, is at least dumb. (1)

The second reply is to invoke our *intuitions* about logicity as the standard against which we judge proposed demarcations. One can find such appeals to intuitions in much work on the demarcation of logic. For example, Sher 1991 writes:

The distinction between logical and extralogical terms is founded on *our pre-theoretical intuitions that logical consequences are distinguished from material consequences in being necessary and formal*. To reject this intuition is to drop the foundation of Tarski’s logic. To accept it is to provide a ground for the division of terms into logical and extralogical. (51, emphasis added)

Feferman objects that Sher’s criterion for logicity assimilates logic too much to mathematics, adding that the persuasiveness of his objection “... will evidently depend on one’s gut feelings about the nature of logic ...” (A:11). Are we reduced, in the end, to weighing one person’s intuitions against another’s gut feelings?

This methodology is ultimately not very satisfying. Our intuitions about logicity are not a kind of perception of an extramental reality: they are historical artifacts, a product

---

<sup>23</sup>I owe the reference to Danielle Macbeth.

of our logical and philosophical *educations*. To the extent that there is intersubjective agreement about them,<sup>24</sup> it should be attributed to a shared tradition, not access to a tradition-independent reality. As Stewart Shapiro and others have pointed out, the very idea of “pretheoretical logical intuitions” is dubious. Students beginning an introductory logic class typically have *inferential* intuitions, but they can be brought to distinguish *logically valid* inferences from materially valid ones only by instruction. All of our intuitions about logic bear the stamp of theory. If we have the intuition that logic must be “formal,” this is not because of some kind of extrasensory perception of the essence of logic, but because we have encountered the idea so often in the course of our philosophical training.

This is not to say that we should *ignore* all of our intuitions about logic. But before invoking these intuitions in justifying or criticizing a proposed demarcation of logic, we ought to seek their sources in the philosophical tradition. By studying the history of conceptions of logic, we can see *why* philosophers have the intuitions they do. Knowing this, we can proceed to ask which intuitions we still have reasons to have, which go together, and which come from incompatible traditions. Historical reflection is a way to make our “brute intuitions” less brute.<sup>25</sup>

It is remarkable how far back one needs to go to achieve the kind of historical understanding I am describing. Twentieth century logicians (e.g., Russell and Tarski) often invoke “formality” as a criterion of logic without saying much about what it means or why it is an appropriate criterion to use in characterizing logic.<sup>26</sup> Responsibility for these tasks is implicitly deferred to a prior (unspecified) tradition. In what follows, I will be arguing that we cannot get clear about the intuitions that guide contemporary debates about the demarcation of logic unless we go all the way back to Kant.

Adapting the Kantian slogan to yet another purpose, we might say: “intellectual history without conceptual analysis may be empty, but analysis without history is blind.”

---

<sup>24</sup>And perhaps there is not much: as Warmbröd 1999 points out, different philosophers have very different intuitions about logic (513).

<sup>25</sup>As Ian Hacking says (in another context): “The ‘fly-bottle’ was shaped by prehistory, and only archaeology can display its shape” (1973:188).

<sup>26</sup>See footnote 21, above.

## 1.5 Prospectus

So much for motivation. The plan for this work is as follows:

In chapter 2, I isolate three “decoy” notions of formality: syntactic formality, schematic formality, and grammatical formality. I call them “decoys” because they distract us from our real target: although they provide us with clear and unproblematic senses in which logic can be said to be “formal,” these senses are not capable of *demarcating* logic (that is, distinguishing it from other disciplines). It is important to hive off these clear and inadequate notions of formality at the beginning, so that we can concentrate on the murkier notions of formality that *might* be capable of demarcating logic.

I turn to these in chapter 3. I isolate three non-equivalent notions of formality that have been invoked (both historically and today) for the demarcation of logic. To say that logic is formal in the first sense (“1-formal”) is to say that it provides constitutive norms for thought as such—a set of rules to which any activity that counts as thought must be held accountable. To say that logic is formal in the second sense (“2-formal”) is to say that it is indifferent to distinguishing features of different objects. And to say that logic is formal in the third sense (“3-formal”) is to say that it abstracts entirely from the semantic *content* of thought. I argue that these three notions of formality are not equivalent; distinguishing them allows us to avoid equivocation and gives us a clearer picture of the conceptual landscape. Moreover, the three notions have very different upshots for the demarcation of logic, as becomes clear when we consider each from the point of view of the projects discussed in section 1.2 (above).

In chapters 4–5, I explore the conceptual archaeology of the three notions of logical formality. Though a full history of logical hylomorphism would be fascinating, I confine myself here to two episodes vital for our self-understanding.<sup>27</sup>

---

<sup>27</sup>In appendix A, I survey logical hylomorphism in ancient and medieval philosophy, with special attention to Abelard’s distinction between perfect and imperfect *inferentia* (the probable ancestor of the later medieval distinction between formal and material consequence). Abelard is particularly interesting, because unlike other medievals, he gives an *argument* for why we must draw a principled distinction between (3-)formally and materially valid inferences. His argument, as I reconstruct it, depends on characteristically medieval ontological assumptions that we (and Kant) would reject.

In chapter 4, I trace modern logical hylomorphism back to Kant. I argue that Kant's logical hylomorphism is an innovation: that none of Kant's modern precursors regarded logic as distinctively "formal." After Kant, on the other hand, the association of logic with form becomes a virtual commonplace: all nineteenth century logical hylomorphism, I argue, derives from Kant. These striking historical facts have a philosophical explanation. When Kant characterizes logic as "formal," he means that it is 3-formal. When he characterizes logic as "general," he means that it is 1-formal. Kant believes that general logic must be formal because, in the context of his transcendental philosophy, 1-formality *implies* 3-formality. Thus Kant's logical hylomorphism is intimately tied up with his transcendental idealism. (Indeed, as I show, both date to about the same period in Kant's thought: 1773-5.) But because of the huge influence of Kant's logical writings on later logic handbooks (first in Germany and later in England), and because the link between Kant's logical hylomorphism and his critical philosophy is not obvious, logical hylomorphism "sticks" in a way that transcendental idealism does not. Faced with the task of finding *something* to mean by the claim that logic is distinctively formal, later writers who reject Kant's overall philosophical framework grasp at 1-formality and 2-formality. In this way, the *word* "formal" comes to be applied to three non-equivalent notions, often without explicit awareness of the differences between them.

In chapter 5, I show how Frege's logicism (in particular, his rejection of the Kantian thesis that objects can be given only through intuition) drives him to reject the Kantian conception of logic. On Frege's mature view, logic is 1-formal, but not 3-formal or 2-formal: like the other sciences, it has its own contentful concepts and relations (and even its own objects), and thus cannot abstract entirely from content. Because Frege rejects crucial tenets of Kant's critical philosophy, Frege can reject the Kantian thesis that in order to be 1-formal, logic must be 3-formal. There is no incompatibility, Frege argues, between taking logic to be a substantive science (i.e., not to be 3-formal) and taking it to provide norms for thought as such (i.e., to be 1-formal). I argue that Frege's conception of logic does not depend on his doctrine (later abandoned by Frege and most contemporary philosophers)

that there are distinctively logical *objects*. The interplay between Kant and Frege on the formality of logic is revealing of the extent to which Kant’s views on logic are intertwined with his larger philosophical views.

In chapter 6, I turn back to the contemporary scene. Whereas historically 3- and 1-formality have been the important notions in discussions of the bounds of logic, it has become common in contemporary philosophy of logic to appeal to *2-formality* in delineating the logical notions.<sup>28</sup> This approach offers the hope of retaining continuity with the tradition of logical hylomorphism without confronting the difficult and obscure notions of 3- and 1- formality. In contrast to these notions, 2-formality is relatively clear, because it can be cashed out in precise mathematical form as the requirement that logical constants be invariant under all permutations of a domain of objects.

I offer a critique of this approach to the demarcation of logic. After giving a precise articulation of the permutation invariance criterion in a general type-theoretic framework, I argue that there are no non-arbitrary grounds for restricting the criterion to classical extensional logics. I show how it can be extended in a natural way to multivalued and intensional logics. When the invariance criterion is presented in its full generality, however, it becomes clear that the application of the criterion—even to classical extensional logics—presupposes a prior articulation of what I call “intrinsic structure” on the basic semantic types.<sup>29</sup> Proponents of the invariance approach simply take for granted that the intrinsic structure on the domain of objects is the “null structure.” Until they offer a justification for this assumption, they have given no reasons for thinking that (for instance) the set-theoretic membership relation is not logical.

At the end of chapter 6, I make a tentative suggestion about how the lacuna in the invariance approach might be filled. I suggest that we can separate “intrinsic” from “non-

---

<sup>28</sup>As we will see, for Kant, 2-formality is not *sufficient* for logicity, and for Frege, it is not even *necessary*.

<sup>29</sup>Some (defeasible) paradigm cases for intrinsic structure: the implication-relevant ordering on multivalued (e.g., the ordering of True and False in two-valued logic), the accessibility relations between possible worlds in modal frames, the structure of “branching histories” on the set of moments in Prior-Thomason style tense logics.

intrinsic” structure on a semantic type by appealing to *1-formality*. Together, 1-formality and 2-formality yield a definite and interesting concept of logical formality. To say that an operator is logical, on my proposal, is to say that it is invariant under all bijections of the basic semantic types that preserve intrinsic structure, where intrinsic structure is the structure on a semantic type that must be mentioned in an account of the relation between the values in that type and the “top-level” semantic notions relevant to correctness of assertion and inference (generally, truth and validity). The proposal imposes a definite task on anyone who wants to claim that a particular logical system is “formal.” Only systems whose concepts and laws can be articulated semantically in terms fundamental to the explanation of correct assertion and inference will count as formal in the sense I have described. I show how this criterion might be used to argue for the logicity of tense logical operators.

In the conclusion (chapter 7), I summarize the main conclusions of the dissertation and make a number of suggestions for applications and future research.

## Chapter 2

# DECOYS

Ask the average philosopher whether logic is formal, and I wager he or she will say “of course.” That is because there are several well-known senses in which logic can be said quite unproblematically to be “formal”:

- Logic can be treated purely syntactically, without reference to the meanings of expressions.
- Logical laws are schematic; that is, they contain “blank spaces” in which any “material” expressions of the appropriate categories can be placed.
- The logical properties of sentences depend only on their *grammatical* forms or structures; that is, on the order and arrangement of their grammatical particles and the grammatical categories of their “categorematic” terms.

What I will argue in this chapter is that these familiar and unproblematic senses of formality cannot do the work of *demarcating* logic—that is, of distinguishing the logical from the non-logical. However, their familiarity and clarity can lull us into thinking that we understand what is meant by claims that logic is *distinctively* formal—i.e., that logic can be *demarcated* by its formality. It is the presence of these “decoy” notions of formality, as I call them, that has prevented logical hylomorphism from getting the critical examination

it deserves. Like decoys in war, they draw attention away from the real target. In this chapter, I aim to reveal them *as* decoys, and thereby make it puzzling again what might be meant by characterizing logic as “formal.”<sup>1</sup>

## 2.1 Syntactic formality

“Formal” often means “syntactic,” that is, “having to do with the symbols themselves, without reference to their meanings.” According to Carnap 1937, for instance,

A theory, a rule, a definition, or the like is to be called formal when no reference is made in it either to the meaning of the symbols (for example, the words) or to the sense of the expressions (e.g. the sentences), but simply and solely to the kinds and orders of the symbols from which the expressions are constructed. (1; cf. 1942:232; 1943:6)

Similarly, Tarski 1965 writes that

... in constructing a deductive theory, we disregard the meaning of the axioms and take into account only their form. It is for this reason that people when referring to those phenomena speak about the purely FORMAL CHARACTER of deductive sciences and of all reasonings within these sciences. (128)

And Gödel 1995 notes that his rules of inference

... are purely formal, i.e., refer only to the outward structure of the formulas, not to their meaning, so that they could be applied by someone who knew nothing about mathematics, or by a machine. (45)<sup>2</sup>

This is the sense in which we talk of “formal systems”—systems definable entirely in terms of effective rules for the manipulation of objects considered as meaningless—and of “formalism” as a philosophy of mathematics. When ambiguity is a danger, I will call this sense of formality *syntactic formality*.<sup>3</sup>

---

<sup>1</sup>Etchemendy 1983, a critique of the view that “[t]wo sentences cannot differ logically if they do not also differ *formally* or *structurally*” (320), considers only the decoy notions of formality discussed in this chapter. The present project was motivated in part by reflecting on whether there might be more to the tradition of calling logic “formal” than the views criticized by Etchemendy.

<sup>2</sup>Unlike Carnap, Gödel does not allow infinitary rules to count as syntactical (ibid., 338), on the grounds that such rules could not “be applied by someone who knew nothing about mathematics.”

<sup>3</sup>A caveat: as Curry 1951 argues, the notion of a formal system is more general than that of the syntax of a language. A formal system is utterly indifferent to the intrinsic natures of its “tokens”

This is a useful and widespread sense of formality, but it is not one that is helpful in delineating logic. The reason is that syntactic formality concerns the *manner* in which a subject matter is treated, not the subject matter itself. Thus, although logic uses formal methods, the formality of these *methods* does not define its *subject matter*. To see why, one need only reflect that syntactic-formal methods can be used in the study of economics, sociology, and geometry, as well as logic. In order to distinguish formal systems of *logic* from formal systems of *geometry*, we need to appeal to something beyond their (syntactic-) formality. What makes a formal system a formal *geometry*—and not, say, a formal economic theory—is that it has some application to the study of space, or at least some structural analogies to other systems that do.

What, then, makes a formal system a formal *logic*? According to Curry, it is the system’s application to the study of “the norms, or principles of valid reasoning” (1963:1). I have suggested in section 1.1 that this answer will not do. But some answer is needed. The capacity for syntactic formal treatment is something logic shares with other disciplines and therefore cannot be used to demarcate logic.<sup>4</sup>

It is sometimes thought, however, that there is *nothing* that stands to syntactic-formal logics as space stands to syntactic-formal geometry, or as matter and energy stand to syntactic-formal physics: that what makes a formal system a *logic* is precisely that there is *no subject matter* to which it is responsible outside of the formal rules themselves. Whereas (primitive terms) or its modes of combining terms into more complex terms (28). Hence it is gratuitous to limit the tokens to linguistic symbols and the modes of combining to “the fundamental operation of concatenation,” as Carnap does. However, for most purposes, the loss of generality in thinking of formality in terms of syntax is trivial, because (as Curry shows) “. . . an arbitrary formal system can be embedded in a syntactical one” (41). Thus the name “syntactic formality” should not mislead, provided that one keeps in mind the essential point that in treating a subject matter syntactic-formally, one abstracts *entirely* from the meanings of (all of) its terms, allowing *only* those claims about them that are explicitly licensed by the rules of the formal system.

<sup>4</sup>Church claims (1956:55) that the requirement that proofs in logistic systems make no reference to any interpretation, proceeding instead by (syntactic) rules alone, can be regarded as “. . . a more precise formulation of the traditional distinction between form and matter. . . and of the principle that the validity of an argument depends only on the form. . . .” But to the extent that the “traditional” distinction was intended to distinguish *logical* or formal validity from non-logical or material validity, Church’s appeal to syntactic formality cannot replace it: syntactic rules can be given to license “material” inferences as well as “formal” ones.

a formal physics can be *applied* to make claims about the world, by relating its fundamental terms to possible observations or physical quantities, a formal *logic* has no *room* for application. It is completely self-contained, in the sense that *every* question one can ask using only the vocabulary of the formal logic is already settled by the rules.<sup>5</sup> On this view, logic is distinguished by the fact that there is nothing to it beyond the syntactic-formal rules—no subject matter to which they might be held responsible.<sup>6</sup> Carnap held this view in his syntactic period:

All of logic including mathematics, considered from the point of view of the total language, is thus no more than an auxiliary calculus for dealing with synthetic statements. *Formal science* has no independent significance, but is an auxiliary component introduced for technical reasons in order to facilitate linguistic transformations in the *factual sciences*. (127) . . . In adjoining the formal sciences to the factual sciences *no new area of subject matter* is introduced. . . . *The formal sciences do not have any objects at all*; they are systems of auxiliary statements without objects and without content. (1934:128)

Can logic can be distinguished from other sciences—even from sciences that can be given a syntactic treatment—by the fact that there is nothing to it beyond what can be syntactically expressed?

No. The putative disanalogy between logic and physics does not hold up. Syntactic-formal logics stand in need of application just as syntactic-formal systems of physics do. For until they are applied, syntactic rules like

Every statement of the form  $S_1 \vee S_2$  is a direct consequence of  $S_1$  (Carnap 1934:126)

are completely useless as “auxiliaries” in scientific inquiry. Until these rules are applied, “is a direct consequence” is merely a syntactic relation, implicitly defined by the rules laid down

---

<sup>5</sup>On Carnap’s 1937 account, for instance, an expression is *logical* just in case it belongs to the largest set of expressions of the language such that every sentence built up from expressions in the set is *determined* as either valid (a consequence of the null set) or contravalid (having every sentence as a consequence) by the rules of the system alone (§50). (By “largest” here I mean maximal: there is no expression that could be added to the set without spoiling the determinacy property. If there is more than one set that meets this description, Carnap instructs us to take their intersection.)

<sup>6</sup>This is not to deny that some syntactic-formal logics might prove to be better tools for certain purposes than others.

for it. Its resemblance to the ordinary notion of *direct consequence* is merely typographical; we could replace it with “is a gabberdegock” without altering anything essential about the rules. By themselves, the syntactic rules do not license us to *infer*  $S_1 \vee S_2$  from  $S_1$ ; they do not tell us that  $S_1$  *implies*  $S_1 \vee S_2$ ; they just tell us that the former “is a direct consequence” of the latter (where we must not assume that the quoted phrase has its English meaning). In order to use the rules as an “auxiliary calculus for dealing with synthetic statements,” we must *apply* them by taking “is a direct consequence” as an inference license or as an indication that a relation of implication holds.<sup>7</sup> In virtue of this application, the system comes to be responsible to a subject matter outside of itself (namely, implication, or correct inference), in precisely the same way that the rules of a formal physics are responsible to extra-syntactic facts about mass, charge, and acceleration.<sup>8</sup>

The fact that every question that can be asked using the logical constants alone is settled by the syntactic rules of the logical system, while the rules of a formal physics leave certain questions open, is irrelevant. To see this more clearly, consider the syntactic rules governing Prior’s (1960) constant “tonk”:

Every statement of the form  $S_1$ -tonk- $S_2$  is a direct consequence of  $S_1$ .

Of every statement of the form  $S_1$ -tonk- $S_2$ ,  $S_2$  is a direct consequence

As Prior points out, introducing these rules into a system trivializes the relation of “direct consequence”: every sentence becomes a “direct consequence” of every other. This result is innocuous enough when “direct consequence” is thought of as a syntactic predicate implicitly defined by the axioms; but as long as “direct consequence” is so considered, the system is useless as an “auxiliary calculus for dealing with synthetic statements.” Once the syntactic system is pressed into scientific service by taking “direct consequence” as an inference license

<sup>7</sup>See Sellars 1953 for a similar observation: “. . . by telling us that transformation rules can be formulated as definitions of ‘direct consequence in S’, Carnap gives the impression that the force of a rule to the effect that expressions of kind A can be ‘transformed’ into expressions of kind B, relates solely to the existence of a structural relationship between these two kinds of expression” (328).

<sup>8</sup>It might be objected that implication and consistency are not “subject matters” in the same sense as mass, charge, and acceleration. Surely there is an intuitive difference in kind here. But characterizing this difference is just part of the task of demarcating logic. We can’t appeal to it in executing that task.

or relation of implication, there emerges a decisive extra-syntactic reason for rejecting the rules: they tell us, wrongly, that every sentence implies every other. Thus the fact that all questions posed using only logical constants can be answered by the syntactic rules alone does not show that logic is not responsible to a subject matter, for there are certain ways of answering these questions (e.g., “does every statement follow from every other?”) that would be *wrong*, given the intended use of the syntactic-formal logics. To say that logic is “completely formal,” then, cannot mean that its syntactic-formal rules are responsible to *nothing* outside of themselves.<sup>9</sup>

In characterizing logic, then, one will eventually need to invoke some semantic concepts, though perhaps only very general ones.<sup>10</sup> The fact that certain logics can be *formulated* in completely syntactic terms, without reference to the meanings of their symbols, does not support the claim that they have no semantic content, as many have thought.<sup>11</sup> Thus the idea that logic is distinctively formal cannot be cashed out in terms of *syntactic formality*. We will have to look elsewhere for a meaning for “formal.”

## 2.2 Schematic formality

It is worth mentioning another notion of formality which gives a kind of *faux* clarity to the idea that logic is distinctively formal. To say that logic is formal is often just to say that it

---

<sup>9</sup>In a critical review of Carnap 1943, in which Carnap attempts to show that the semantic properties of the logical constants can be determined entirely by syntactic rules, Church 1944 objects that Carnap’s systems cannot be said to fix the meanings of the logical connectives *by syntactic means alone*, because they do so only under the assumption that the syntactic notions of *junctive* and *involution* “...receive a particular kind of interpretation” (497). Even Carnap’s requirement that “...the syntactical relation of derivability receive a particular kind of interpretation” seems “arbitrary” (497), given his project. Why is it okay to stipulate a semantic interpretation for *this* syntactic relation, but not (say) for the sign “ $\vee$ ”? This is essentially the same point I am making.

<sup>10</sup>The point I am making does not hinge on taking *truth* as the central semantic notion, as in Hacking 1979. It applies equally to approaches that take inference or provability as central, as in Prawitz 1978.

<sup>11</sup>Cf. Cresswell 1973: “One reason why ‘truths of logic’ have been historically thought of in the narrow sense we have called ‘logical validity’ is undoubtedly that the logically valid principles which result from treating only a small number of symbols as constants allow of reasonably simple formulation. It is even possible to formulate them without any reference to interpretation at all, a fact which is no doubt partly responsible for the idea that they are true independently of any content” (34).

concerns itself with inference *patterns* or *schemata* (“forms”) whose instances are all correct inferences, no matter what the instantiating “matter”: for example,

All As are Bs.  
All Bs are Cs.  
Therefore, all As are Cs.

Particular inferences are called “formally valid” or even “valid in virtue of their forms” when they are instances of schemata all of whose instances are correct inferences. When ambiguity threatens, I will call this notion *schematic formality*.

Schematic formality has played an important role in the history of logic. That logic is schematically formal is a presupposition of the use of formal counterexamples to establish logical invalidity—a technique that goes back to Aristotle’s *Prior Analytics*. It is also the guiding idea behind the modern semantic (or model-theoretic) definition of logical consequence, due to Tarski 1936. In the paper that gives the first statement of the semantic definition, Tarski takes schematic formality as a criterion of adequacy for the success of his account. He takes the following condition (called “F” for “formal”) to be a necessary condition for a sentence  $X$  to be a logical consequence of a class  $K$  of sentences:

(F) If, in the sentences of the class  $K$  and in the sentence  $X$ , the constants—apart from purely logical constants—are replaced by any other constants (like signs being everywhere replaced by like signs), and if we denote the class of sentences thus obtained from  $K$  by ‘ $K'$ ’, and the sentence obtained from  $X$  by ‘ $X'$ ’, then the sentence  $X'$  must be true provided only that all sentences of the class  $K'$  are true. (415)

The definition of logical consequence as the preservation of truth (or some other desirable semantic feature<sup>12</sup>) in all models—that is, in all semantic reinterpretations of the language’s non-logical expressions—ensures that logical consequence will be schematic-formal in the sense of condition (F).

---

<sup>12</sup>Note that the generalization over interpretations is separable from the appeal to *truth preservation* in Tarski’s definition. The generalization over interpretations is a common feature of semantic definitions of consequence for all kinds of logics (relevance logics, intuitionistic logics, modal logics, etc.); it has nothing to do with classical logic in particular.

The problem with schematic formality is that it contains two lacunae, which must be filled before it can yield any definite verdicts on the formality of particular inferences. To fill the first lacuna, one must specify which features of inferences are part of the *patterns* and which are replaceable or schematic elements. To fill the second, one must specify the range of expressions that can replace various types of schematic elements in these patterns. Verdicts of formality can differ wildly depending on how these lacunae are filled. For example, even the inference

This substance turns litmus paper red.  
Therefore, this substance is an acid.

would count as formally valid relative to the pattern:

A turns litmus paper red;  
Therefore, A is an acid.

where A takes singular terms as instances. On the other hand, the paradigmatically logical consequence

All cats are mammals;  
All mammals are animals;  
Therefore, all cats are animals.

would *fail* to be formally valid if “all” were not taken as part of the inference schema.

To see the importance of the second lacuna, notice that the inference

All cats are animals;  
Therefore, some animals are cats.

could be either formally valid or not relative to the schema

All As are Bs;  
Therefore, some Bs are As.

depending on whether or not terms with empty extensions are allowed as instances of the schematic letters A and B.<sup>13</sup>

---

<sup>13</sup>For the importance of the second lacuna, see Evans 1976, section II.

The same two lacunae are present in the Tarskian (model-theoretic) account of logical consequence. This account must be completed with (i) a definition of truth in an interpretation and (ii) a specification of the class of interpretations. These tasks are usually discharged by giving (i) a specification of the *logical constants* and their (fixed) interpretations, and (ii) a specification of the *semantic categories* for the non-logical expressions. As Tarski realized, nonstandard choices for the logical constants (i) yield odd relations of “logical consequence:”

Underlying our whole construction is the division of all terms of the language discussed into logical and extralogical. This division is certainly not quite arbitrary. If, for example, we were to include among the extralogical signs the implication sign, or the universal quantifier, then our definition of the concept of consequence would lead to results which obviously contradict ordinary usage. (1936:418)

Etchemendy 1990 has reminded us that the output of the model-theoretic definition of consequence is similarly affected by assumptions about the semantic categories (ii). He points out, for instance, that if the class of semantic values for singular terms (i.e., the domain) were finite, there would be no countermodel to the sentence “ $[\forall x\forall y\forall z(Rxy \ \& \ Ryz \supset Rxz) \ \& \ \forall x\sim Rxx] \supset \exists x\forall y\sim Ryx$ ” (“if R is a transitive, irreflexive relation, then R has a minimal element”) (118).

These lacunae could of course be filled by stipulation. But the resulting notion of formality would have little interest from the point of view of demarcating logic. *Every* inference is schematic-formal relative to *some* pattern.

Alternatively, instead of filling the lacunae, we might leave them open and define a *relational* notion of schematic formality: *formal relative to constants C and semantic categories S*.<sup>14</sup> The difficulty lies in finding a *use* for this relational concept—in saying what *turns on*

<sup>14</sup>Tarski’s own views on these matters seem to have fluctuated. At the end of his 1936 article, he notes that unless “important objective arguments” can be found that “justify the traditional boundary between logical and extralogical expressions,” we will be forced to take logical consequence as a *relative* concept, “. . . which must, on each occasion, be related to a definite, although in greater or less degree arbitrary, division of terms into logical and extralogical” (420). In a 1944 letter to Morton White, he seems to plump emphatically for the latter option: “Let us not forget even for a minute that any definition of ‘logical term’ and ‘logical truth’ can only be given in terms of a

an inference's being formal relative to a particular C and S. At any rate, the philosophical projects canvassed in section 1.2 (above) all seem to presuppose a non-relational concept of logicity. It would be of little interest for a logicist to point out that the laws of arithmetic are formal relative to the constants 0, S, +, -,  $\div$ ,  $\times$ , =, &,  $\sim$ ,  $\forall$  and the semantic categories *natural number* and *relation on natural numbers*. Such a result would not show why arithmetic is knowable *a priori* or generally applicable. Nor would a relational notion of formality be of any use in defining the boundary between pure mathematics and other disciplines (unless that boundary exhibits the same kind of relativity). Nor, finally, would a relational notion of formality help to explain how logic can serve as a framework for the explicitation of scientific and mathematical concepts and theories. The challenge for someone who defines a relational notion of formality, then, is to articulate its philosophical significance.<sup>15</sup>

A third alternative would be to supplement schematic formality with a principled delineation of the logical constants and an account of the semantic categories. *This* would yield a notion of formality by which we might delineate logic. But schematic formality would only be the skeleton, as it were, of this notion: the muscles and organs (that is, the hard work) would be contained in the accounts of the logical constants and semantic categories. The hard questions about the bounds of logic would be questions about what it is for an expression to be a logical constant or what the semantic categories ought to be. And the notion of formality to which some philosophers have appealed in addressing *these* questions—for instance, in rejecting an expression as a logical constant on the grounds that it has “material content”—is plainly not schematic formality. It would be a mistake, then, to think that by invoking schematic formality, one has given an account of the sense in determined language (or to a determined class of languages)” (29). But in his 1966 lecture, he presents a principled demarcation of the logical notions in terms of permutation invariance.

<sup>15</sup>Brandom 1994 countenances notions of K-formality corresponding to *any* set K of constants: on his account, inferences can be “. . . good in virtue of their zoological, moral, or theological form” (105). In the limiting case, where K includes *all* of the language's simple expressions, *any* good inference will be good in virtue of its K-form. However, Brandom does not give a philosophical *use* for any of these concepts of formality, and they do not play a theoretically significant role in his system. For scepticism about the significance of the relational notions, see Etchemendy 1990:100-106.

which logic is distinctively formal. At most, schematic formality is one piece of the story. It cannot help us to answer the hard questions about the bounds of logic: it cannot tell us, for instance, whether tense logics or set theories should count as logics.<sup>16</sup>

## 2.3 Grammatical formality

It is sometimes thought that the lacunae in schematic formality can be filled by appealing to the *grammar* of a language, so that logical form becomes grammatical form or structure. For example, Quine 1980 writes:

It is a general practice, in intellectual pursuits, to argue from the truth of one sentence to the truth of another. Some such arguments are the business of logic, others not. They belong to logic if they hinge purely on the structure of the sentences concerned, rather than depending on content. But the structure of sentences consists in grammatical constructions. Here, then, is the intimate connection between grammar, truth, and logic. Logic studies the truth conditions that hinge solely on grammatical constructions. (17)

Quine proposes that we identify the logical terms of a language with its grammatical *particles*—the expressions associated with the modes of construction by means of which complex sentences are built up recursively from simpler ones in the grammarian’s reconstruction of a language’s syntax—and that we look to the grammar’s *lexicon* for categories and ranges of the schematic letters.<sup>17</sup> Similarly, Dummett 1973 proposes that logical expressions of a language are the expressions by means of which complex sentences are built up, step by step, from atomic ones; while non-logical expressions are the simple expressions that make up atomic sentences (213).

Quine’s and Dummett’s proposals might be regarded as spelling out precisely how logic depends on form or structure—*grammatical* structure. However, there are three serious

---

<sup>16</sup>Here I am merely echoing a point made by Bolzano 1837. After suggesting schematic formality as a philosophically hygienic way to cash out the claim that logic is formal (§12), Bolzano warns that this sense of formality cannot delineate logic from other disciplines: “. . . in this sense [of ‘form’], there are infinitely many divisions which concern only the form of propositions; but logic has to lift out only those which yield something useful; it has to make us acquainted only with the sorts or kinds of propositions which have something peculiar in their scientific treatment” (§186, my translation).

<sup>17</sup>For discussion, see Føllesdal 1980 and Harman 1984.

problems for proposals of this kind. Our examination of these problems should make it clear that logic cannot be demarcated by reference to *grammatical formality*.

### 2.3.1 The immanence of grammar

Since only a language has a grammar, grammatical formality is unabashedly relational: *formal relative to a language*. That's okay: it might be that the line between logic and non-logic is language-relative. The problem is that, as Quine acknowledges, there are in general many acceptable ways of formulating the grammar for a given language—that is, of recursively defining the infinite set of its meaningful sentences: “[o]ne and the same language, indeed—one and the same infinite set of sentences, anyway—can of course be generated by different constructions from different lexical beginnings” (1986:60).<sup>18</sup> In particular, the distinction between lexicon and particle is *immanent*, not transcendent: it has a role within a particular regimentation of grammar, but there is no account of it that floats free of this role, no extra-theoretical “fact of the matter” about how it should be drawn. All that is given is the set of meaningful expressions of the language. How we generate these recursively is up to us.

To use an example from Quine 1980, we might regard the comparative suffix “-er” as either a particle or as an integral part of lexicon items. On the first approach, comparative adjectives like “taller” would not be items in the lexicon; they would be generated from non-comparative adjectives (like “tall”) by a comparative-forming operation that concatenates the particle “-er.” On the second approach, “taller,” “bigger,” and so on would just be items in the lexicon. The two competing grammatical analyses will yield different separations of consequences into “formal” and “non-formal.” On the first approach, the inference

Taller(Joe, Jack)  
 Taller(Wilma, Joe)  
 Therefore, Taller(Wilma, Jack)

will count as a formally valid inference, since it cannot be turned into a bad (non-truth-

---

<sup>18</sup>I am presuming here that languages are not individuated by their grammars. On Quine’s view, a language is simply a set of meaningful expressions.

preserving) inference by uniform substitution of lexicon items for lexicon items in the same categories. But on the second approach, the same inference will not count as formally valid, since it could be turned into a bad inference by substituting “In-love-with()” for “Taller()” (now considered as a lexical item).

More radically, we could choose to regard the two-place truth-functional connectives as members of a small lexical category instead of as particles (cf. Quine 1986:28-9). Instead of recognizing four grammatical operations that form one sentence from two sentences (one that takes P and Q and yields  $P \wedge V \wedge Q$ , one that takes P and Q and yields  $P \wedge \& \wedge Q$ , and so on, where ‘ $\wedge$ ’ is the sign for concatenation), we would recognize a single grammatical operation that forms one sentence from two sentences and one connective (taking P, Q, and ‘&’ to  $P \wedge \& \wedge Q$ , for example).

If grammatical formality is to be used as a criterion for the demarcation of logic from non-logic, then, it must be combined with some way of privileging one grammar out of the plurality of adequate ones. Quine’s approach is to privilege the grammar that is best suited for a canonical language for science; that is, the one that allows the most economical representation of the truth conditions of sentences. He proposes that “what we call logical form is what grammatical form becomes when grammar is revised so as to make for efficient general methods of exploring the interdependence of sentences in respect of their truth values” (1980:21). By “efficient” and “economical,” Quine seems to mean “with a minimum of structure”: on this basis he argues that the “-er” in “taller” should not be taken as a particle (which would make the transitivity of “taller” a logical truth), since the same truth conditions can be captured by taking “taller” as a member of the lexicon and adding explicit premises to license the transitivity inferences (1980:19-21, 1986:77-8). A pragmatic maxim guides judgments of efficiency and economy: “expose no more structure than is needed for the purpose at hand” (22). The privileged grammar, then, is the one that makes semantic relations perspicuous using a minimum of grammatical structure. Indeed, Quine’s proposal might best be conceived not as a criterion for logical constancy but as a pragmatic principle for language design: design your canonical languages so that grammatical structure

reveals semantic relations, but do so with as little grammatical structure as possible. It is a good principle, for many purposes. We *like* to use languages in which grammatical structure reveals semantic relations. But this does not show any interesting sense in which logic might be separated from other disciplines on the grounds of its “formality.”

We are looking for a sense of formality that might be capable of demarcating logic. A language-relative notion might fit the bill. But a notion that depends on how we *look at* a language—that is, how we regiment its grammar—will not. The distinction between logic and non-logic may be relative to a language, but it should not be relative to a way of *describing* a language, on pain of losing all theoretical significance. Quine can accept a grammatical criterion for the logical constants only because he does not think that the line between logic and non-logic *has* any deep theoretical significance:

We may recall ... that the very distinction between grammar and lexicon is immanent, and admits of alternative adjustments even in the analysis of one and the same language. As this distinction varies, the distinction between logical truth and other truth varies with it. Insofar, the demarcation of logical truth rests on the whim of the descriptive grammarian. (1986:96)

One might reasonably take this consequence as a *reductio ad absurdum* of Quine’s criterion of demarcation.<sup>19</sup>

### 2.3.2 Identity

The second problem with the grammatical criterion is that “=” is usually taken to be a logical constant, even though there is no *grammatical* basis for distinguishing it from non-logical relation terms, like “is taller than,” “gives presents to,” or “surrounds.” Since “=” is

<sup>19</sup>In this discussion, I have assumed (as Quine does) that the only non-pragmatic constraint on the grammar of a language is its ability to finitely represent the set of the language’s meaningful expressions. This assumption is justified, I think, for artificial languages, but it may not be justified for natural languages. Linguistics and cognitive science may give good grounds for preferring one of two grammars that are equivalent in the set of meaningful expressions they generate, on the grounds that it corresponds more closely with the psychological mechanisms responsible for speakers’ linguistic competence. However, it is doubtful that the psychologically “real” grammar would yield a notion of formality useable for demarcating logic, and I am not aware of any recent proposals along these lines. In the early 1970s, there was a proposal (“generative semantics”) to meld grammatical form with logical form by letting the goals of each constrain the other (see Massey 1975:74-5). But this program seems to have self-destructed by the late 1970s.

intersubstitutable *salva congruitate* with an indefinitely large class of such terms, it ought to be treated as a member of the lexicon, not a particle (see Quine 1986:22, 28). This leaves us with a choice: abandon the claim that “=” is a logical constant, or abandon the grammatical demarcation of logic.

Quine takes the first route. Though he acknowledges several good reasons for taking identity to be a logical notion (62), he is unwilling to give up his grammatical demarcation of logic. Instead, he reaches a compromise, observing that we can have identity theory without taking “=” as a logical constant, provided that the language has a determinate, finite stock of predicates. The trick is to *define* (or “simulate”) “=” by laying down axioms guaranteeing substitutability of identicals (1986:63). “On this plan,” Quine remarks, “the identity sign does not qualify as a grammatical particle, but its laws still belong to logic” (1980:28).

This approach has its costs. It means that “=” must be redefined every time the language is expanded. And although Quine’s technique preserves the *deductive* strength of identity theory, it does not preserve its *model-theoretic* power. For example, it is easy to construct a sentence of pure identity theory that is true only in domains containing exactly 2 objects:

$$\text{(Two)} (\exists x)(\exists y)(\sim x=y \ \& \ (\forall z)(z=x \vee z=y))$$

But this is no longer possible when “=” is treated in Quine’s way, as a defined non-logical term. To see why, suppose the language contains just two predicates (both one-place), F and G. Then “=” is defined as follows:

$$\text{(IdDef)} \ x = y \iff (Fx \equiv Fy \ \& \ Gx \equiv Gy)$$

Let T be the theory consisting of (Two) and (IdDef), with “=”, “F”, and “G” the only non-logical constants. Now consider the following model, M:

Domain:	{1, 2, 3, 4}
Extension of “F”:	{1, 2}
Extension of “G”:	{3, 4}

Extension of “=”:  $\{ \langle 1, 1 \rangle, \langle 1, 2 \rangle, \langle 2, 1 \rangle, \langle 2, 2 \rangle, \langle 3, 3 \rangle, \langle 3, 4 \rangle, \langle 4, 3 \rangle, \langle 4, 4 \rangle \}$

One can easily see that  $M$  is a model of  $T$ , even though it has a domain with four objects. Thus Quine’s compromise sacrifices some of the characterization strength of identity theory.

Dummett rides the other horn of the dilemma, opting, in effect, to qualify the grammatical demarcation. In a footnote (1973:22), he proposes that identity should count as logical on the grounds that it allows us to express “quantifier conditions” (properties of predicates that are invariant under permutations of the elements of the domain) that we could not express otherwise. This is a semantic condition, not a grammatical one, and admitting it seriously compromises the explanatory force of the grammatical criterion. For suppose we are in doubt as to the logicity of some constant—say, the “ $\in$ ” of set theory—that functions grammatically as a relational expression. We cannot confidently exclude it from the bounds of logic on the basis of the grammatical criterion, once we have allowed that there can be exceptions to the criterion on well-motivated semantic grounds. The grammatical criterion is reduced to a “rough and ready” guideline, and thereby loses much of its theoretical interest. It cannot be the whole story about logicity.

### 2.3.3 Grammatical chauvinism

Even if we forget about identity and the immanence of grammar, the grammatical criterion stands guilty of what I call “grammatical chauvinism.” That is, it takes what is merely a contingent feature of our most familiar logical systems and presents it as a general rule.

Call a grammatical category *uniform* if either all of its members are logical constants or all are non-logical constants. Otherwise, call the category *mixed*. Call a *language* uniform just in case *all* of its grammatical categories are uniform.<sup>20</sup> It is *almost* true that all standard logical languages (standard propositional logics, standard first- and second-order logics, standard modal and tense logics) are uniform. (Almost, because of the problem of identity—but forget about that for now.)

<sup>20</sup>I assume that the grammar of the language is fixed; here we abstract from the difficulties discussed in section 2.3.1, above.

Grammatical demarcations of logic take this fact (or almost-fact) about the languages with which we are most familiar and make it into a general rule:

**Uniformity thesis.** Every language must be uniform.

I will argue that the uniformity thesis is false: there is nothing wrong with languages that are not uniform (or almost-uniform). They are just unfamiliar. But the grammatical criterion for logical constanthood, which sorts expressions into logical and non-logical on the basis of their grammatical functions, plainly *implies* the uniformity thesis. So if the uniformity thesis is false, we have reason to reject the grammatical criterion for logical constanthood.

To see that the uniformity thesis is false, one has merely to design a language that is not uniform. Start with the language LP of standard first-order logic, with logical constants “ $\forall$ ”, “ $\exists$ ”, “ $\&$ ”, “ $\sim$ ”, “ $\vee$ ”, “ $\equiv$ ”, “ $\supset$ ”. These particles fall into the following grammatical categories:

- One-place variable-binding connectives: “ $\forall$ ”, “ $\exists$ ”
- One-place sentential connectives: “ $\sim$ ”
- Two-place sentential connectives: “ $\&$ ”, “ $\vee$ ”, “ $\equiv$ ”, “ $\supset$ ”

LP is clearly uniform: all of the logical constants fall into one of these three categories, and none of these categories contains any non-logical expressions. Now add a one-place variable-binding connective “C”, interpreted as “there is at least one cat such that . . . .”<sup>21</sup> Call the resulting language LP+C. The sensible thing to say about LP+C, I think, is that it is not uniform: “C” is a non-logical constant in the same grammatical category as a logical constant.

A true believer in the grammatical criterion might resist this conclusion in one of two ways, neither of them savory. The first is to deny that LP+C is a possible (or coherent, or legitimate) language, perhaps on the grounds that there could not be a lexically *simple*

<sup>21</sup>Strictly speaking, we ought to say: “introduce a grammatical mode of construction that takes any open or closed sentence P and any variable u and yields the sentence “(Cu)P”. “C” is merely the particle that marks this particular mode of combination.

term with the meaning of “C”. I’m not sure what would support such a claim, so I will pass it by.<sup>22</sup>

The second way is to take “C” to be a logical constant. That would save the grammatical criterion, but at what cost? One would have to say that the sentence

$$(Cx)(Cy)(\sim x=y), \text{ i.e., “there are at least two cats”}$$

is a sentence of pure logic, while

$$(\exists x)(\exists y)(\sim x=y \ \& \ \text{Cat}(x) \ \& \ \text{Cat}(y))$$

which translates the same way into English, contains non-logical vocabulary. It is difficult to believe that such a notion of logicity could have any interesting use.

Rather than taking such desperate measures to salvage the uniformity thesis and the grammatical criterion for logical constanhood, it seems healthier to reject both and to disconnect questions of logicity from questions of grammatical role. We should acknowledge the possibility of languages with non-logical expressions in the grammatical roles traditionally occupied only by logical constants. Such languages have been explored: Kuhn 1981 shows us how we might construct an “operator logic” with non-logical sentential connectives. I wholly concur with Kuhn’s motivation: “[t]he question of whether an expression is logical or not is a philosophical question, and its answer should not be forced on us by a formal apparatus” (495).

Summing up: there are three problems with invoking *grammatical* structure to explain the distinctive “formality” of logic. First, there is the problem of multiple grammars: in general, the same set of meaningful expressions can be generated recursively in many different ways. We must either bring in non-grammatical considerations to choose between them or accept the consequence that the bounds of logic depend, in part, on “the whim of the descriptive grammarian.” Second, there is the problem of “=”, which is usually taken

---

<sup>22</sup>One might argue that *grasp* of the meaning of “C” presupposes grasp of some simpler (non-logical) expression, like “cat.” No doubt it does: but now we are doing semantics, not grammar. Dummett, at any rate, is clear that the notion of simplicity relevant for his criterion is simplicity of *expressions*, not simplicity of *senses* (24–5).

to be a logical constant, even though grammatically it is indistinguishable from other two-place predicates. Finally, there is the problem of grammatical chauvinism: the grammatical demarcation of logic takes what is merely a contingent fact about standard logical languages (their uniformity) and enshrines it as a general criterion.

## 2.4 Conclusion

We have surveyed three notions of formality. Although they are all clear and relatively unproblematic, none is capable of demarcating logic from other disciplines. Syntactic formality concerns a way in which logic (as well as many other disciplines) might be *treated*; it does not characterize the *subject matter* of logic. Schematic formality gives no verdicts about logicity until its two lacunae are filled. Grammatical formality founders on the problems of immanence, identity, and chauvinism.

In calling these notions “decoys,” I do not mean to suggest that they are unintelligible, or even that they can play no role in demarcating logic. The problem with syntactic, schematic, and grammatical formality is that they give us (at best) only *part* of the story. They do not give us an understanding of what is *distinctive* about logic. What they do give us is a false impression that we know what it means to say that logic is distinctively formal. But in order to understand this idea, we must turn away from these clear but inadequate notions of formality and accustom our eyes to the considerably murkier notions of formality by which logic has historically been delineated.

## Chapter 3

# THREE NOTIONS OF LOGICAL FORMALITY

Logic, it is often said, is distinctively *formal*: it concerns itself with those relations of implication and consistency that turn only on the “forms” or “formal features” of thoughts or statements, abstracting from the “matter” or “content.” This kind of talk is so common as to be nearly invisible. Asked what it means, philosophers will often appeal to one or more of the senses of formality canvassed in the previous chapter. As we have seen, however, these senses are decoys. Although they are clear and philosophically innocuous, they are not fit for the task of carving out the province of logic. Logic can be treated syntactically—but so can other disciplines. Logical laws are schematic—but so are non-logical laws. Logical validity turns on grammatical form—but only if we set up the grammar to reveal and systematize logical validity. What, then, *is* meant by “formal” when philosophers use it to delineate logic (e.g., in the passages quoted in section 1.3.3, above)?

In this chapter, I will distinguish three different notions of logical formality. All three have played a role in the literature since Kant, and all three continue to shape our intuitions about logic. I make no claim to completeness: there may be other notions of logical formality. I am confident, however, that these are the main notions in play:

To say that logic is **1-formal** is to say that its norms are *constitutive* of concept use *as such* (as opposed to a particular kind of concept use). 1-formal laws are the norms to which any conceptual activity—asserting, inferring, supposing, judging, and so on—must be held responsible.

To say that logic is **2-formal** is to say that its characteristic notions and laws are indifferent to the particular identities of different objects. 2-formal notions and laws treat each object the same (whether it is a cow, a peach, a shadow, or a number). Mathematically, 2-formality can be spelled out as invariance under all permutations of the domain of objects.

To say that logic is **3-formal** is to say that it abstracts entirely from the semantic content or “matter” of concepts—that it considers thought in abstraction from its relation to the world and is therefore entirely free of substantive pre-suppositions.

We can get at these three notions by construing “formal” as “independent of content or subject matter.” What does it mean to say that logic is independent of content or subject matter? That depends on what we mean by “content” or “subject matter.”

1. We might mean “particular domain of application.” In that case, to say that logic abstracts from content or subject matter is to say that it is applicable without qualification, in any domain—that it is normative for thought or concept use *as such*, or 1-formal.
2. We might mean “particular object or individual.” In that case, to say that logic abstracts from content or subject matter is to say that it pays no heed to distinguishing features of individuals, but treats them all the same—i.e., that it is 2-formal.
3. We might mean “semantic content.” In that case, to say that logic abstracts from content or subject matter is to say that it abstracts entirely from the semantic contents of concepts, claims, and inferences—i.e., that it is 3-formal.

In the next three sections (3.1– 3.3), I flesh out these three notions of formality in more detail, with examples. Then, in section 3.4, I show that the three notions of formality are conceptually independent of each other, and in section 3.5, I show how they provide three ways of spelling out the “generality” or “topic-neutrality” of logic. Finally, in section 3.6, I consider the significance of each of the notions for logicism.

### 3.1 1-formality

Consider two kinds of norms for playing chess. On the one hand, there are prescriptions for playing chess well, for carrying out the Queen’s Indian defense, or for playing an endgame with two rooks. These are norms that are applicable only in specific situations, or only given certain goals or interests: they are “hypothetical,” in the sense of Kant’s “hypothetical imperatives.” On the other hand, there are the rules of the game. These norms apply to chess playing *as such*, because they are *constitutive* of chess playing. One might violate them (intentionally or inadvertently) and still count as playing chess. One might even be ignorant of some of them and still count as playing chess. But unless these norms are *binding* on one’s moves, one is not playing chess, but some other game. They are “categorical” norms for chess.

We can make the same distinction among norms for *thought* (which, following Kant, I will take as a blanket term for all concept use, including judging, inferring, supposing, asserting, and even perceiving, insofar as it involves the application of concepts). On the one hand, there will be norms for certain kinds of concept use (asserting, laying down legal decisions), or for the use of certain kinds of concepts (physical concepts, chemical concepts, moral concepts). These will be “hypothetical” norms for thought, norms with limited or conditional application. On the other hand, there may be norms for concept use *as such*: norms that are *constitutive* of concept use. If there are any such norms, they will be “categorical” norms for thought, and they will be universally and unconditionally applicable.

Kant appeals to this distinction in distinguishing the laws of formal logic from the laws of the special sciences. The laws of the special sciences are *contingent* laws of the understanding: laws “...without which a certain determinate use of the understanding would not occur” (JL:12). They are contingent not in the sense that they could have been otherwise (among them Kant includes the laws of geometry), but in the sense that they are only *conditionally* applicable to thought: “...it is contingent whether I think of this or

that object, to which these particular rules relate” (JL:12).<sup>1</sup> Thus, for example, the laws of physics are norms of thought *about matter and energy*. They are “contingent” in the sense that one can think without being constrained by them—provided one does not think about matter and energy. To interpret an activity as thought about matter and energy is to hold it subject to evaluation by these laws.

The laws of logic, by contrast, are defined as the “. . . necessary laws of the understanding and of reason in general, or what is one and the same, of the mere form of thought as such . . .” (JL:13). By “necessary laws of the understanding,” Kant means “. . . those [laws] without which no use of the understanding would be possible at all. . .” (JL:12), that is, the norms *constitutive* of thought. Similarly, in the first *Critique* he says that general logic “. . . contains the absolutely necessary rules of thought without which there can be no employment whatsoever of the understanding” (KrV:A52/B76). In my terminology, this is a demarcation of logic by its *1-formality*.

As I will show in chapter 4, Kant himself does not use “formal” to mean *1-formal* (instead, he uses “general”), but many later writers do. This use of “formal” is common in nineteenth century works on logic, especially those influenced by Kant. For example, Hyslop writes:

. . . Logic is a science of the *formal* laws of thought. They are the laws which are not only essential to it, but which are the same whatever the subject-matter involved in our reasoning. The laws of thought remain the same in the reasonings of Astronomy, Physics, Politics, or Ethics, but the ‘matter’ changes and does not affect the validity of the process. The ‘form’ of our reasoning in all these cases is essential to its being such a process. Hence Logic, as a science, is ‘formal,’ and deals only with the ‘formal’ principles of thought in distinction from the material objects of reason. (Hyslop 1892:12)

Frege, who also distinguishes logic from the special sciences by its 1-formality, usually uses

---

<sup>1</sup>Kant’s use of “necessary” and “contingent” in JL:12 must be distinguished from his use of the same words in JL:14. In JL:12, necessary and contingent laws of the understanding are two different brands of *norms*. The difference is that contingent laws are not applicable to thought as such, but only to thought about some particular objects. In JL:14, contingent laws are *non-normative* psychological laws of “how we do think,” while necessary laws are norms governing “how we ought to think.” The distinction of JL:12 corresponds to the first *Critique*’s distinction between general and special logics (A52/B76), while the distinction of JL:14 corresponds to the distinction between pure and applied logic (A53/B77).

“general” to indicate this feature.<sup>2</sup> But in at least one passage, he uses “formal” in the sense of *1-formal*:

...the basic propositions on which arithmetic is based cannot apply merely to a limited area whose peculiarities they express in the way in which the axioms of geometry express the peculiarities of what is spatial; rather, these basic propositions must extend to everything that can be thought. And surely we are justified in ascribing such extremely general propositions to logic. I shall now deduce several conclusions from this logical *or formal* nature of arithmetic. (Frege FTA:95, emphasis added)<sup>3</sup>

It is important to be clear about the sense in which logical laws are, on this tradition, normative for thought as such. The point is *not* that nothing can count as thought unless it *conforms to* the laws of logic. That would make logical error impossible, and it would make nonsense of Kant’s claim that logic is a *normative* discipline, since the way we *ought* to think would turn out to be the only way we *could* think. When Kant says that we cannot think except “according to” the laws of logic (JL:12), he means that our thought must be *responsible to* the laws of logic for its assessment. Just as the throwing of a baseball does not count as a *pitch* unless it is liable to assessment in light of the rules of baseball, so no cognitive activity counts as *thought* unless it is liable to assessment in light of the laws of logic. And just as there can be an illegal pitch, so there can be an illogical thought. What makes it a thought is not that it *conforms to* the laws of logic, but that the laws of logic are normative for it. To say that the laws of logic are norms for thought as such, then, is not to say that it is impossible to think illogically, but only that it is impossible to think illogically and be thinking *correctly*.<sup>4</sup>

Moreover, as every teenage driver soon discovers, one can be normatively constrained by laws one does not even *acknowledge*. The laws to which thought about matter and energy

<sup>2</sup>The laws of logic, he says, “. . . are the most general laws, which prescribe universally the way in which one ought to think if one is to think at all” (GGZ:xv).

<sup>3</sup>Joan Weiner picks up this language in her 1986 article: “[Frege’s logic]. . . is meant to be nothing more nor less than a means of setting out the formal rules of all thought, the rules which make the use of the understanding possible.” Curiously, in the corresponding discussion in her 1990 book, she replaces “formal rules of all thought” with “true rules of all thought” (79). As we will see in chapter 5, Frege too gives up this hylomorphic language in his later writing, as he gets clearer about his differences with Kant.

<sup>4</sup>This, I take it, is the main upshot of Frege’s “logical aliens” thought experiment (GGZ:xvi-xvii).

is responsible are the *true* laws of physics, whether we know them are not. The same point applies to the laws of logic. Thus, to say that they are norms for thought *as such* is not yet to say that we know them *a priori*. It is true that Kant and Frege held that logical laws had both these characteristics, but this is a substantive thesis. The two characteristics may go together, but let us not conflate them.

On the other hand, it *does* follow from logic's being normative for thought as such that it is *necessary* in a strong sense. We can *think* about possible worlds or situations in which the physical laws of the actual world do not hold—"where animals speak and stars stand still, where men are turned to stone and trees turn into men, where the drowning haul themselves up out of swamps by their own topknots. . ." (Frege FA:§14). Such thought violates certain contingent norms for thought—norms for correct thought about animals, stars, forces, and so on—and thus does not count as correct thought about animals, stars, forces, and so on. But considered merely as *thought*, it cannot be faulted, because it does not violate the norms for thought as such. What makes it possible for us to think *correctly* about such counterfactual worlds, then, is that we can prescind from *some* of the norms of thought—by acknowledging our thought as counterfactual, as concerning mere possibility—while continuing to acknowledge others. In the limiting case of *logical* possibility, we prescind from *all* contingent norms of thought, acknowledging only the norms for thought *as such*. But if we try to prescind from these norms, too—say, by thinking about a possible world in which contradictions are true—then *no* norms remain to which our concept use can be held responsible. In this case it is no longer recognizable as concept use at all, since concept use is *essentially* evaluable as correct or incorrect. We can correctly think about what the world would be like if the laws of physics were different, but not about what it would be like if the laws of logic were different. This is the sense in which the norms for thought *as such* are necessary: it is impossible to think at all, even counterfactually, without being constrained by them.

In addition to necessity, 1-formality carries with it another feature that has been important in the philosophy of mathematics: independence from intuition or sensibility. Kant saw

the dependence of mathematics on intuition as a restriction on its domain of applicability: mathematical norms were binding only on reasoning about objects capable of being given to the senses (and famously led to antinomies outside this realm). In this framework, to show that some mathematical laws are 1-formal—that is, normative for thought *as such*—would be to show that they apply to objects independently of whether these objects can be given in intuition, and thus that they do not depend on experience or pure intuition for their *justification*. Frege argued backwards along this path, from the applicability of arithmetic outside the realm of the intuitable to its 1-formality or logicity (FTA:94-5, FA:§14).

It is worth emphasizing that the claim that there *are* norms for concept use *as such* is nontrivial. It would take argument to rule out the possibility that *all* norms for concept use are contingent in the sense of Kant’s JL:12. Even if all thought must be subject to norms, it is not obvious that there must be any one norm to which all thought is subject. It is not obvious that, as Frege says, “[t]hought is in essentials the same everywhere: it is not true that there are different kinds of laws of thought to suit the different kinds of objects thought about” (FA: iii). It might be that thought is *compartmentalized*, so that there are norms for thought about different kinds of objects, but no global norms for thought as such. To the extent that it rules out this possibility, the thesis that logic is 1-formal is hardly trivial.

### 3.2 2-formality

To say that logic is formal is to say that it does not concern itself with specific content. One way this might be cashed out—the way we have just explored—is to say that logical norms are applicable to thought about any subject matter whatever—to thought *qua* thought. But there is also another approach, one that eschews talk of norms and domains of applicability. This approach starts from the idea that what makes content *specific* is its concern with *particular individuals*. It is clear that the concept *horse*, the relation *is taller than*, and the quantifier *every animal* have specific content, because they all distinguish between Lucky Feet, on the one hand, and the Statue of Liberty, on the other:

- “Lucky Feet is a horse” is true; “The Statue of Liberty is a horse” is false.
- “The Statue of Liberty is taller than Lucky Feet” is true; “Lucky Feet is taller than the Statue of Liberty” is false.
- The truth of “Every animal is healthy” depends on whether Lucky Feet is healthy, but not on whether the Statue of Liberty is healthy.

On the other hand, the concept *is a thing*, the relation *is identical with*, and the quantifier *everything* do not distinguish between Lucky Feet and the Statue of Liberty. In fact, they do not distinguish between *any* two particular objects. As far as they are concerned, one object is as good as another and might just as well be switched with it. Notions with this kind of indifference to the particular identities of objects might reasonably be said to abstract from specific content—to be “formal.” Tarski seems to have something like this in mind when he says:

...since we are concerned here with the concept of logical, i.e., *formal*, consequence, and thus with a relation which is to be uniquely determined by the form of the sentences between which it holds, this relation cannot be influenced in any way by empirical knowledge, and in particular by knowledge of the objects to which the sentence X or the sentences of the class K refer. The consequence relation *cannot be affected by replacing the designations of the objects referred to in these sentences by the designations of any other objects.* (Tarski 1936:414-5, emphasis added)

Call this notion of formality “2-formality.” I will argue in chapters 4 and 5 that 2-formality did not historically play a significant role in the demarcation of logic. It has become important as a way of spelling out the distinctive “formality” of logic only in the twentieth century, probably because it is so much clearer than 1- or 3-formality. What is particularly appealing is that the notion of *indifference* by which 2-formality is defined can be cashed out in a mathematically precise way as invariance under a group of transformations. The technique goes back to Felix Klein, who proposed it in 1872 as a way of defining different geometries, and it is helpful to see first how it applies to geometry.

Klein observed that we can distinguish the various geometries by the groups of transformations under which their characteristic notions are invariant. Consider first the familiar case of Euclidean geometry. Its characteristic notions—*parallel*, *similar*, *congruent*, etc.—are indifferent to the absolute spatial location and orientation of the figures to which they apply. If a Euclidean sentence is true of a particular figure on a plane, it will remain true no matter how we move the figure across the plane, rotate it, proportionally stretch or shrink it, or reflect it across a line. That is, Euclidean notions are invariant under the group of similarity transformations. Similarly, the notions of affine geometry are invariant under the group of affine transformations (which preserve straightness of lines, but not angle size). The fact that affine notions are invariant under more transformations than Euclidean notions captures the fact that they are indifferent to more features of figures than the Euclidean notions. For example, affine geometry does not distinguish between different kinds of triangles (equilateral, isosceles, scalene). We can take this process even further. Topology, which is invariant under the group of bicontinuous transformations (transformations that preserve connectedness, but not the straightness of lines), is indifferent to the differences between squares and circles, treating both as simple closed curves.

A number of philosophers have suggested that Klein's method of demarcating geometries by their invariance under different groups of transformations can be extended to logic. One can capture the indifference of logical notions to the particular identities of objects by demanding that they be invariant under the group of *all* transformations of the objects in a domain. Proposals of this kind can be found in Mautner 1946, Mostowski 1957:13, Tarski 1966, Scott 1970:161, Dummett 1973:22 n, McCarthy 1981, van Benthem 1989, Sher 1991 and 1996, McGee 1996, and Shapiro 1998:99.<sup>5</sup>

For present purposes, this approach can best be illustrated through examples.<sup>6</sup> Unary first-order quantifiers can be modeled semantically as functions from sets (predicate exten-

---

<sup>5</sup>Tarski seems to have arrived at his proposal independently of a similar proposal by Mautner 1946. The work on permutation invariance in McCarthy 1981 and Sher 1991 was, in turn, independent of Tarski, whose lecture was not published until 1986. (Sher reports being pleasantly surprised to find her proposal seconded by such a luminary.)

<sup>6</sup>Technical definitions and a full discussion can be found in chapter 6, below.

sions) to truth values. Now consider the quantifier “all chickens,” which takes every set that contains all the chickens in the domain to True and every other set to False. Suppose the domain of objects includes two chickens and two cows. Let  $K$  be a set containing two chickens, so that “all chickens” takes the value True on  $K$ . If we permute the objects in the domain by switching the chickens for the cows, so that  $K$  is now a set containing two cows, then “all chickens” takes the value False on  $K$ . Since the truth of sentences containing “all chickens” is sensitive to permutations of the domain, “all chickens” is not permutation-invariant. On the other hand, the numerical quantifier “at least three things,” which takes every set with three or more members to True and every other set to False, *is* permutation-invariant. No matter how the elements of a set  $K$  are permuted, it will always contain the same number of members, so no permutation can affect the value of “at least three things” on  $K$ .

Permutation invariance can be regarded as a precise technical gloss on the idea of indifference to the particular identities of objects. Whereas “all chickens” is sensitive to the difference between chickens and cows, permutation-invariant notions—like quantifiers and identity—are insensitive to the particular identities of the objects to which they apply: “there are at least three Fs” will be true whether the Fs are numbers, people, places, or diamonds, provided there are at least three of them. The same is not true of the proprietary notions of arithmetic and set theory: addition, for instance, is sensitive to the *differences* between particular numbers.

Which notions are permutation-invariant? The notions we regard as basic to (extensional) logic—universal and existential quantifiers, identity, and (trivially) the truth functions—are all permutation-invariant. But there are also a few others, most prominently the cardinality quantifiers: “there exist at least  $\alpha$  things such that . . .,” and “there exist exactly  $\alpha$  things such that . . .,” for every cardinal  $\alpha$ . Where  $\alpha$  is finite, these quantifiers are already definable in standard first-order logic. But the addition of cardinality quantifiers with infinite  $\alpha$  yields a significant expansion of the expressive power of first-order logic (see Tharp 1975). Indeed, as Feferman (A) shows (drawing on McGee 1996), the addition of all the first-order permutation-invariant quantifiers to a language gives it the power of full

second-order logic.

We motivated 2-formality as a way of thinking about abstraction from *specific* content. Thus the claim that logic is 2-formal leaves open the possibility that logic is concerned with *general* content and very general facts about the world. 3-formality, to which we now turn, precludes *any* concern with content.

### 3.3 3-formality

To say that logic is 3-formal is to say that it abstracts entirely from the semantic content of thoughts (or interpreted sentences). The word “entirely” is essential here: a logic that abstracts from the contents of *some* concepts (“specific” ones, like “horse” or “red”), but not from the contents of others (“general” or logical ones, like existence, identity, and conjunction), does not count as 3-formal.

What is left to consider when we abstract entirely from the semantic content of a thought? Some philosophers would say “nothing.” But some have held that there is more to thought than the concepts it employs: there is also the way they are put together. On Kant’s view, for example, “. . . all judgments are functions of unity among our representations” (KrV:A69/B93-4), and the various modes of unity determine the possible *forms* of judgments (A70/B95):

The *matter* of the judgment consists in the given representations that are combined in the unity of consciousness in the judgment, the *form* in the determination of the way that the various representations belong, as such, to one consciousness. (JL:101)

General logic, on Kant’s view, abstracts from the matter and considers only the form of judgments: it considers only the way in which concepts are united in judgments.<sup>7</sup>

General logic. . . abstracts from all content of knowledge, that is, from all relation of knowledge to the object, and considers only the logical form in the relation

---

<sup>7</sup>This is a bit of a simplification. Kant applies the matter/form dichotomy at three levels—concepts, judgments, inferences—and general logic concerns itself with the formal element at each level (see JL:§2, §18, §59).

of any knowledge to other knowledge; that is, it treats of the form of thought in general. (Kant KrV:A55/B79)

For example, general logic treats “all horses are mammals” simply as the unification of two concepts in a universal, affirmative, categorical, and assertoric judgment. It abstracts entirely from the content of the concepts. The way in which the concepts are united in thought is not, for Kant, a further constituent of the thought (a “binding” concept), but a feature of the thought’s *form*.

As a result of this abstraction from content—that is, from the relation of thought to its objects<sup>8</sup>—logic cannot yield any knowledge about the world or any real truths:

...since the mere form of knowledge, however completely it may be in agreement with logical laws, is far from being sufficient to determine the material (objective) truth of knowledge, no one can venture with the help of logic alone to judge regarding objects, or to make any assertion. (A60/B85)

But this limitation of 3-formal logic is also an advantage. Precisely because logic abstracts from all relation to objects in the world, there is no substantive question regarding its *adequacy* to these objects, as there is in the case of mathematics:

That logic should have been thus successful is an advantage which it owes entirely to its limitations, whereby it is justified in abstracting—indeed, it is under obligation to do so—from all objects of knowledge and their differences, leaving the understanding nothing to deal with save itself and its form. But for reason to enter on the sure path of science is, of course, much more difficult, since it has to deal not with itself alone but also with objects. (B ix)

Kant wastes no time explaining the possibility of our *a priori* knowledge of the laws of logic: he takes this to be unproblematic (JL:15). Indeed, his explanation of our *a priori* knowledge of mathematics *presupposes* that the understanding has transparent knowledge of its own forms, considered independently of their relation to objects. Kant’s Copernican turn—the doctrine that “. . . we can know *a priori* of things only what we ourselves put into them” (KrV:B xviii)—could not help to explain our *a priori* knowledge of mathematics if

---

<sup>8</sup>Since for Kant the content of a concept depends on its relation to an object (A69/B94, A139/B178, A155/B194-5, A239/B298, JL:§2), these formulations are equivalent (A59/B83, A63/B87).

it were problematic how we could have *a priori* knowledge of “what we ourselves put into” things.

The claim that logic is 3-formal presupposes that we can distinguish between the constituents and the form of thought, in such a way that the latter can be understood apart from the former. Kant’s idealist successors reject 3-formality because they reject this distinction.<sup>9</sup> But many twentieth century philosophers accept the distinction, demarcate logic by its concern with the *form* of thought or conceptual inquiry, and draw the same consequences as Kant:

1. that logic alone tells us nothing about the world;
2. that the world does not *constrain* logic, so that logical knowledge and knowledge about the world are fundamentally different; and
3. that “logical truths,” if there are such things, do not state facts.

I will give three examples.

### 3.3.1 Schlick and Einstein

Schlick and other early logical empiricists with backgrounds in neo-Kantianism invoke a sharp form/content distinction to explain how pure geometry can be an *a priori* science without invoking Kant’s “pure forms of intuition.” In the absence of pure forms of intuition, Schlick reasons, the only way to avoid making pure geometry empirical is to disconnect it entirely from the real world. He does this by regarding geometric concepts as *implicitly defined* in terms of their logical relations to each other, as specified by a system of axioms in which only the logical constants are interpreted. Implicitly defined concepts, Schlick says, “have no association or connection with reality at all. . .” (1925:37, quoted in Coffa

<sup>9</sup>Hegel writes: “Logic is usually said to be concerned with forms *only* and to derive the material for them from elsewhere. But this ‘only,’ which assumes that the logical thoughts are nothing in comparison with the rest of the contents, is not the word to use about forms which are the absolutely real ground of everything” (1827:49, 51). Cf. Bradley 1883:519-24.

1991:176).<sup>10</sup> Einstein articulates this view lucidly:

As far as the laws of mathematics refer to reality, they are not certain; and as far as they are certain, they do not refer to reality. It seems to me that complete clearness as to this state of things first became common property through that new departure in mathematics which is known by the name of mathematical logic or ‘Axiomatics.’ The progress achieved by axiomatics consists in its having neatly separated the logical-formal from its objective or intuitive content; according to axiomatics the logical-formal alone forms the subject-matter of mathematics, which is not concerned with the intuitive or other content associated with the logical-formal. . . . [On this view it is clear that] mathematics as such cannot predicate anything about perceptual objects or real objects. In axiomatic geometry the words ‘point,’ ‘straight line,’ etc., stand only for empty conceptual schemata. (Einstein 1921:28, 31)

But this way of accounting for our *a priori* knowledge of pure geometry presupposes that the “logical-formal” notions in terms of which geometrical concepts are implicitly defined are themselves devoid of all “objective or intuitive content.” In other words, it presupposes that logic is 3-formal. For if logic had objective content, then the implicitly defined pure geometrical concepts would have objective content too, and pure geometry would “refer to reality.” It is not simply because these concepts are *implicitly defined* that Schlick and Einstein can take them to be “logical-formal,” but because they are implicitly defined in terms of their *logical* relations to each other. Concepts implicitly defined in a *non-logical* background language would have the same kind of objective content as the resources used to define them.<sup>11</sup>

### 3.3.2 Carnap

Carnap’s influential distinction between formal and factual sciences is motivated by similar concerns about the problem of mathematical knowledge:

In this distinction we [the Vienna Circle] had seen the way out of the difficulty which had prevented the older empiricism from giving a satisfactory account of the nature of knowledge in logic and mathematics. . . . Our solution, based on

---

<sup>10</sup>See Friedman 1990 for an interesting discussion of the difficulties into which Schlick is led by his strict form/content distinction.

<sup>11</sup>Cf. section 1.2.2, above.

Wittgenstein's conception, consisted in asserting the thesis of empiricism only for factual truth. By contrast, the truths in logic and mathematics are not in need of confirmation by observations, because they do not state anything about the world of facts, they hold for any possible combination of facts. (Schilpp 1963:64)<sup>12</sup>

Note that if this solution is to be non-trivial, "factual truth" must not simply *mean* "truth about the empirical world." Carnap is therefore committed to the view that the true sentences of logic and pure mathematics do not state facts of any kind. They "...do not express any matters of fact, actual or nonactual" (Carnap 1934:126). Echoing Kant's characterization of logic, Carnap says: "*The formal sciences do not have any objects at all; they are systems of auxiliary statements without objects and without content*" (128). But precisely *because* they do not have content, they are not constrained by facts about the world. Instead, they define the "linguistic framework" in terms of which such facts can be stated. This linguistic framework can be regarded as the *way* in which contentful concepts are constructed and related to one another: the *form* of thought, or as Carnap would prefer to put it, of scientific inquiry.

Carnap diverges from Kant in maintaining that the form of thought is, in some sense, a matter of *choice* or *convention*, and that consequently there can be many equally legitimate logical frameworks (Schilpp 1963:64). Conventionalism was Carnap's answer to a question that naturally arises for the view that logic is 3-formal: if logic is not constrained by the *world*, what is the source of its objectivity? Carnap's answer has drawn heavy fire (e.g., from Quine 1963 and Prior 1976:123-4). But here I am primarily interested in showing *that* Carnap holds logic to be 3-formal, not *how* he articulates this position or whether he succeeds in defending it.

---

<sup>12</sup>Carnap acknowledges the influence of Wittgenstein's *Tractatus* here, and Wittgenstein does seem to be advocating a version of the view that logic is 3-formal: "The propositions of logic are tautologies.... The propositions of logic therefore say nothing.... Theories which make a proposition of logic appear substantial are always false" (1922:§§6.1, 6.11, 6.111). For a sophisticated account of the relation between Carnap's view of logic and Wittgenstein's, see Friedman 1988.

### 3.3.3 Nagel

Ernest Nagel defends a similar view in his article “Logic Without Ontology” (1956:ch. 4). His goal is “to make plausible the view that the role of the logico-mathematical disciplines in inquiry can be clarified without requiring the invention of a hypostatic subject matter for them. . .” (57). Like Carnap, he draws a sharp distinction between logical principles—“principles whose function it is to institute a desired order into inquiry”—and “statements about the explicit subject matter of inquiry” (72). Thus, for example, the principle of non-contradiction is not a claim about things in the world, but rather a criterion for the use of the logical terminology “same respect,” “same attribute,” “belong,” and “not belong” (58):<sup>13</sup>

Accordingly, the interpretation of the principle as an ontological truth neglects its function as a *norm* or *regulative principle* for introducing distinctions and for instituting appropriate linguistic usage. To maintain that the principle is *descriptive* of the structure of antecedently determinate ‘facts’ or ‘attributes’ is to convert the *outcome* of employing the principle into a *condition* of its employment. (60)

Nagel revealingly characterizes the view he is *opposing* as the view that logical laws “. . . are not formal or empty” and that “. . . they tell us something about the *actual* world” (66). We may infer that the view he is defending is a version of the claim that logic is 3-formal. Logical laws are implicit definitions that define the framework or *form* of rational inquiry about the world (80). Because they do not represent constraints on the world, they can be known *a priori*; but for the same reason, they can tell us nothing about the world.

## 3.4 Independence of the three notions

These, then, are three things that might be meant by calling logic “formal”: first, that it provides norms constitutive of concept use as such; second, that it is indifferent to the particular identities of individuals; third, that it abstracts entirely from the semantic content of the concepts used. It is useful to think of these three notions of formality in relation to

---

<sup>13</sup>Similar ideas lie at the root of C. I. Lewis’s neo-Kantian pragmatism (1929:245-6).

three Kantian dualisms identified in Brandom 1994:614-18: 1-formality can be understood in terms of the dualism of thought and sensibility, 2-formality in terms of the dualism of general and singular, and 3-formality in terms of the dualism of structure and content (see figure 3.1).

Figure 3.1: Three notions of logical formality.

Notion	Description	Dualism
1-formality	normativity for thought <i>as such</i>	thought/sensibility
2-formality	indifference to particular identity	general/singular
3-formality	abstraction from semantic content	structure/content

In certain philosophical frameworks, these three dualisms line up neatly, and logic can be said to be formal in all three senses. For example, in Kant's transcendental idealism, sensibility provides the *content* for cognition and is the source of all *singular* representations, whereas structure and generality depend on thought. It should not be surprising, then, that Kant holds logic to be formal in all three senses. (The relation between Kant's philosophy of logic and his transcendental idealism will be discussed in more depth in chapter 4.)

But the three notions need not line up in this way. For example, Frege rejects the Kantian connection between sensibility and content. Thus, although like Kant he holds that logic is 1-formal, he rejects the Kantian idea that logic is 3-formal. (Frege's rejection of 3-formality will be discussed in chapter 5.) On Frege's view, logic is about the world in just the same sense as physics, only its concepts are more general. Frege also severs the Kantian connection between sensibility and singularity: for Frege, thought has its own objects (extensions and truth values). Thus, Frege's logic (unlike Kant's) is not 2-formal: it respects differences between particular objects and employs concepts that are not permutation-invariant. 1-formality implies neither 2-formality nor 3-formality.

Conversely, 2-formality does not imply 1-formality. There may be general truths expressed in permutation-invariant vocabulary that are not binding on thought *as such*. For example, a finitist might claim that (as a matter of contingent fact) the world contains no

more than  $n$  objects.<sup>14</sup> Since “there are no more than  $n\dots$ ” is a permutation-invariant quantifier, the finitist will have to count this claim as 2-formally true. But she need not count it as 1-formally true: it is coherent for a finitist to suppose that the number of objects in the world could have been larger than it is.

Indeed, Kant holds that arithmetic and algebra are 2-formal, but not 1-formal or 3-formal.<sup>15</sup> These sciences do not, on Kant’s view, have their own objects; rather, they are directed toward “...objects without distinction” (A63/B88), from the particular features of which they abstract completely (A717/B745). But although the characteristic concepts of arithmetic and algebra do not respect individual differences between objects, the laws of arithmetic and algebra are not normative for thought *as such*; they are binding only on thought about objects capable of being exhibited in sense. Nor do they abstract entirely from semantic content: they do not abstract from the concept of *magnitude*, a concept that (in Kant’s view) can be given content only through construction in pure intuition.

3-formality does not imply 1-formality, either. On the positivists’ view of logic, for instance, the necessity of logic is simply a reflection of the rules for the use of a particular language (cf. Ayer 1946:77). Thus logical laws are not normative for thought as such, but only for thought in a particular framework. If we do not want to be bound by them, we can simply use another language. In this respect, they are analogous to Kant’s “contingent laws of the understanding,” which are only binding on us if we intend our thought to relate to certain objects.

Finally, 2-formality and 3-formality can come apart. It is important not to overlook the difference between abstracting from all relation to objects—that is, from all content of concepts—and abstracting from all *differences* between objects. As we have seen in articulating the notion of 2-formality, it is not necessary to abstract from content entirely

---

<sup>14</sup>Some philosophers may question the very intelligibility of such a claim, on the grounds that it makes sense to talk of the number of objects only relative to a sortal concept. But the finitist’s claim can be understood as: “there are no more than  $n$  objects of all sorts.” Such a claim would be true if there were  $m$  *basic* sortal concepts ( $S_1\dots S_m$ ) such that (a) the sum of the number of  $S_i$ ’s,  $1\leq i\leq m$ , is less than or equal to  $n$ , and (b) for every sortal concept  $S$ , every  $S$  (i.e., everything that is an  $S$ ) is identical with some  $S_i$  for some  $1\leq i\leq m$ .

<sup>15</sup>See section 4.1.2, below.

in order to abstract from *specific* content. Many contemporary advocates of the 2-formality of logic (e.g., Sher, Shapiro) would repudiate the view that logic is 3-formal.<sup>16</sup> On the other hand, the Carnapian brand of 3-formality allows 3-formal logic (broadly construed) to contain singular terms (such as “2” and “4”) and non-permutation invariant function terms (such as “+”) (1934:124). For example, Carnap 1934 describes a language in which “ $2+2=4$ ” is an “auxiliary statement” with “no factual content” (126). Such a language would be 3-formal but not 2-formal. It appears, then, that neither 2-formality nor 3-formality entails the other.

If the three notions of formality are equivalent in certain philosophical contexts, then, that is because of special features of those contexts. The notions are conceptually independent. As long as they are not explicitly distinguished, however, there is always danger of confusion and equivocation. For example, advocates of the permutation invariance criterion for logical constants often seem to conflate distinct senses of formality in motivating their proposals and connecting them with traditional conceptions of logic. According to van Benthem 1989, “[t]he traditional idea that logical constants are not concerned with real content may be interpreted as saying that they should be *preserved* under those operations on models that change content, while leaving general structure intact” (317). But interpreting historical talk of the formality of logic as inchoate talk of permutation invariance would be a distortion. Historically, 2-formality did not play an important role in debates over the bounds of logic: the “traditional idea” was that logic is 1-formal or 3-formal. And van Benthem has done nothing to connect his proposal with *these* senses of formality or “indifference to real content.” The ambiguity of “formal” allows him to dodge the question of what permutation invariance or 2-formality has to do with *logicality*, as traditionally conceived. Similar unclarity leads Sher to suggest that her account of the logical constants as isomorphism-invariant notions answers Russell’s difficulty about “what is meant by saying

<sup>16</sup>For the claim that logic is 2-formal, see Sher 1996:672-8, Shapiro 1998b:99. Sher and Shapiro do not explicitly repudiate 3-formality, because they do not draw the distinctions I have drawn. But they both reject “foundationalist” approaches to logic: approaches that take logic to have the kind of epistemic priority over contentful mathematics that it would have were it 3-formal (Sher 1996:680, Shapiro 1991:ch. 2).

that a proposition is ‘true in virtue of its form’ (1996:683-4). The first step to securing the kind of historical continuity Sher and van Benthem are seeking is to distinguish the different notions of “formality” that are in play.

### 3.5 Formality, generality, and topic-neutrality

Discussions of the demarcation of logic do not always make heavy weather of “formality.” They commonly appeal to the *generality* or *topic-neutrality* of logic instead.<sup>17</sup> In this section, I will argue that these notions are trifurcated in precisely the same way as formality. By the “generality” (or “topic-neutrality”) of logic, I will show, one might mean either 1-formality, 2-formality, or 3-formality. The upshot is that this dissertation has something to contribute to a whole range of work on the demarcation of logic, not just work that invokes the word “formal.”

What does it mean to say that logic is maximally general, or topic-neutral? We might start with this suggestion:

**(Gen-1)** To say that logic is general or topic-neutral is to say that it is not about anything in particular.

But the notion of *aboutness* is notoriously hard to pin down. We can see some of the difficulties if we consider set theory and arithmetic. On the one hand, they seem to have special subject matters (sets and numbers, respectively). On the other hand, they can be used in discourse on virtually any topic.<sup>18</sup> Suppose one understood only the following words in a paragraph: “set,” “is a member of,” “five,” and “fewer than.” One could confidently say that it was about sets and numbers, yet in a broader sense one would have no idea what it was about. *Any objects* can be considered as a set or numbered. So are set theory and

<sup>17</sup>The term “topic-neutral” is due to Ryle 1954:116. For the claim that logic is characterized by its topic-neutrality, see also Peacocke 1976:229, Benacerraf 1973:411, McCarthy 1981:504, Sainsbury 1991:314, Strawson 1952:41, Haack 1978:5-6, Wright 1983:133. The last three of these connect topic-neutrality with “formality,” as does Ryle. Schroeder-Heister 1984:104 implies that a constant that is not topic-neutral has “specific material content.”

<sup>18</sup>Indeed, *set* is Gödel’s prime example of a formal or “universally applicable” concept (Wang 1996:267).

arithmetic topic-neutral or not? Are they about nothing in particular, or are they about sets and numbers (cf. Boolos 1975:517)?

Goodman 1961 defines a technical notion of “absolute aboutness” that would give definite answers to these and other questions about aboutness. On Goodman’s account, a statement  $S$  is *absolutely about* an object  $k$  just in case there is a statement  $T$  such that . . .

- (i)  $T$  contains an expression  $e$  designating  $k$ , and
  - (ii)  $T$  follows logically from  $S$ ,
  - (iii) no generalization of  $T$  with respect to  $e$  (or any part of  $e$ ) follows logically from  $S$ .
- (253)<sup>19</sup>

It follows from this definition (together with Goodman’s tacit assumption that the logical constants do not designate anything) that logical truths are not absolutely about anything (256) and that logically equivalent statements are absolutely about the same things (258). Goodman’s notion is useless for our purposes, however, because clauses (ii) and (iii) of the definition presuppose that it is settled what counts as logic (253-4). Thus, if set theory is part of logic, set-theoretic truths will not be about anything; while if set theory is not part of logic, set-theoretic truths will be about sets and classes of sets. It would be circular, then, to appeal to Goodman’s criterion in an account of logicality.

We might try something simpler:

**(Gen-2)** To say that logic is general or topic-neutral is to say that its truths and inference rules do not mention any particular objects, i.e., contain no singular terms essentially.<sup>20</sup>

But this is pretty obviously inadequate. The sentence “ $(\forall x)(x \text{ is a cat} \supset x \text{ is a mammal})$ ” does not contain any singular terms, but it is a truth of biology, not of logic (cf. Carnap 1942:232). So the generality of logic cannot consist in the fact that logical truths contain no

<sup>19</sup>I have reformulated Goodman’s definition slightly, without altering its content.

<sup>20</sup>The point of the qualification “essentially” is to allow logical truths like “Ben is tall or Ben is not tall.” The singular term “Ben” does not occur in this sentence essentially, because we could substitute any other singular term without changing the truth value of the sentence.

singular terms essentially. The use of particular *general terms* can spoil maximal generality as surely as the use of particular singular terms.

The most obvious fix is to forbid mention of particular properties, functions, and relations, as well as particular objects. This is Russell's approach from 1913 on:

A proposition which mentions any definite entity, *whether universal or particular*, is not logical: no one definite entity, of any sort or kind, is ever a constituent of any truly logical proposition. (1913:97-8, emphasis added)

Certain characteristics of the subject [mathematics/logic] are clear. To begin with, we do not, in this subject, deal with particular things *or particular properties*: we deal formally with what can be said about *any* thing or *any* property. . . . It is not open to us, as pure mathematicians or logicians, to mention anything at all, because, if we do so, we introduce something irrelevant and not formal. (1920:196-7, emphasis added)

Putting Russell's proposal into a less ontologically committed form, we get:

**(Gen-3)** To say that logic is general or topic-neutral is to say that its truths and inference rules contain no singular terms, predicates, relation terms, or function terms essentially.

However, (Gen-3) will only work if the logical expressions do not count as predicates, relation terms, or function terms. “=” is especially problematic, because it is in the same grammatical category as “is taller than” (see section 2.3.2, above). Doesn't it signify a particular relation, namely, identity? And don't logical truths like “ $\forall x(x=x)$ ” mention this relation?

One might bite the bullet and take “=” to be non-logical. But a similar question arises for the quantifiers and sentential operators. We do not usually think of these as predicates, relation terms, or function terms (though Frege thought of quantifiers as higher-order predicates, and sentential operators as ordinary function terms). So the essential presence of quantifiers and sentential operators in logical truths does not compromise their generality, according to (Gen-3). But suppose our language contains intuitively *non-logical* quantifiers or sentential operators, like “(Cx)” (for all cats . . .) or “(Kn<sub>C</sub>)” (Bill Clinton knows that . . .).<sup>21</sup> (Gen-3) will blindly count truths containing these terms essentially as

<sup>21</sup>See section 2.3.3, above.

“general” or “topic-neutral,” and that seems wrong. But how could we modify (Gen-3) to treat these terms differently from the standard quantifiers and truth-functional operators, without begging the question?

Russell 1913 solves this problem by denying that the logical expressions function semantically in the same way as non-logical expressions of the same grammatical categories. They do not signify constituents of propositions at all; instead, they indicate *form*:

“Logical constants”, which might seem to be entities occurring in logical propositions, are really concerned with pure *form*, and are not actually constituents of the propositions in the verbal expression of which their names occur. (1913:98; cp. 1920:199)

It is because Russell treats logical constants in this way that he can take (Gen-3) as his account of logical generality. Thus generality, for Russell, amounts to a version of 3-formality.<sup>22</sup>

It should be clear, however, that *Russell's* way of understanding the generality of logic is not for everyone. Many philosophers have held that logic is maximally general or topic-neutral without holding that it is 3-formal. Those who do not see a semantic difference in kind between “ $(\forall x)$ ” and “ $(Cx)$ ” will not want to cash out the generality of logic in the way that Russell does in 1913. And even those who hold that logic is 3-formal may want to mean something else by the claim that logic is general.<sup>23</sup>

A natural suggestion is to spell out the generality or topic-neutrality of logic in terms of the generality or topic-neutrality of the expressions that occur essentially in logical truths or inference rules—the “logical constants.”

**(Gen-4)** To say that logic is general or topic-neutral is to say that its truths or inference rules contain only general or topic-neutral expressions essentially.<sup>24</sup>

<sup>22</sup>Russell's theory of “logical forms” is notoriously problematic. See Griffin 1993:177, Hylton 1984:389-91.

<sup>23</sup>In chapter 4, I will argue that Kant falls into this category.

<sup>24</sup>This is Russell's solution from 1903-1911: he characterizes a “formal” proposition as a “proposition which does not contain any other constant than logical constants,” where the logical constants are “purely formal concepts” (1911:288).

But it turns out to be remarkably difficult to say what it is for an *expression* to be general or topic-neutral.

Let us start with Ryle's criterion for topic-neutrality:

We may call English expressions 'topic-neutral' if a foreigner who understood them, but only them, could get no clue at all from an English paragraph containing them what that paragraph was about. (1954:116)

The problem with this criterion is that one might answer the question "what is this paragraph about?" at many different levels of generality. Consider the following paragraph:

If *Bob Wills* had not *recorded* the *tune*, *someone* else certainly would have. Many *Texas fiddlers* were *playing* it at the time. Though it might not have become *famous*, it could not have *escaped being put on vinyl*.

This paragraph is about:

- (1) "Faded Love"
- (2) a particular tune recorded by Bob Wills
- (3) Bob Wills and Texas fiddle music
- (4) music
- (5) persons and things
- (6) the past
- (7) historical possibilities
- (8) objects

Even the original passage, out of context, does not tell us enough to give answer (1). If we delete all the proper names,

If ...had not *recorded* the *tune*, *someone* else certainly would have. Many ...*fiddlers* were *playing* it at the time. Though it might not have become *famous*, it could not have *escaped being put on vinyl*,

we cannot give answers (1)-(3), but we can still give (4)-(8). If we now delete all the underlined words,

If ...had not ...the ..., some ... else certainly would have. Many ... were ... it at the time. Though it might not have become ..., it could not have ...,

we cannot give answers (1)-(5), but we can still give answers (6)-(8). We still have *some* clue about what the passage is about: we know, for instance, that it is not a geometrical proof. Finally, if we delete everything but

If ... not ... the ..., some ... else .... Many ... it .... Though it ... not ..., it ... not ...,

we can still give answer (8)—as we could not, for instance, for the sentence “To be red is to be colored.” *Every* expression (except perhaps “if” and the like) gives us *some* information about the topic of a paragraph. Instead of a firm dividing line between topic-neutral and topic-specific expressions, then, Ryle’s criterion gives us a spectrum of varying degrees.

Perhaps this is all we should ask for. Lycan 1989 suggests that it is a mistake to look for a line between logical and non-logical expressions, since topic-neutrality comes in degrees: truth-functional expressions are more topic-neutral than quantifiers, which are more topic-neutral than tense and modal operators, which are more topic-neutral than epistemic expressions, and so on.<sup>25</sup> Yet it is widely held that the generality of the paradigm logical constants is different in kind, and not just in degree, from that of “quickly,” “red,” or “person.” Is there any way to cash this distinction out in a principled and motivated way?

One place we might look for a sharp dividing line is the permutation invariance criterion, discussed in section 3.2, above. The permutation invariance criterion might be thought to give a precise sense in which logical notions (the semantic values of logical constants) are “not about anything in particular” (cf. (Gen-1), above):

**(Gen-5)** To say that logic is general or topic-neutral is to say that its fundamental notions are invariant under all permutations (or bijections) of the domain of objects.

Not only does the permutation invariance criterion settle the borderline cases for which the topic-neutrality criterion could give no clear verdict, it also provides a kind of *explanation*

---

<sup>25</sup>It would probably be more accurate to say that the relation *is more topic-neutral than* is a partial order: the modal operators and the tense operators, for instance, represent two different spectra of relative topic-neutrality, not two (comparable) positions on a single spectrum.

for the fact that logical expressions can be used in discourse about any topic:

Logic, on the present conception, takes certain general laws of formal structure and, using the machinery of logical terms, turns them into general laws of reasoning, applicable in any field of discourse. The fact that biological, physical, psychological, historical, . . . structures obey the general laws of formal structure explains the generality (“topic neutrality”) of logic: some references to formal structure (to complements and unions of properties, identity of individuals, non-emptiness of extensions, etc.) is interwoven in all discourse, and therefore logic (the logic of negation and disjunction, identity, existential quantification, etc.) is universally applicable. (Sher 1996:674-5)

Thus 2-formality provides one way of understanding the generality of logic, just as 3-formality provides another. But some philosophers think of logic as general in neither of these senses. For example, Frege says that “. . . the task we assign logic is only that of saying what holds with the utmost generality for all thinking, whatever its subject matter,” and that “. . . logic is the science of the most general laws of truth” (PW:128). But he does not think that logic is 3-formal (see chapter 5, below). Kit Fine 1998 suggests that permutation invariance (2-formality) “. . . is the formal counterpart to Frege’s idea of the generality of logic” (1998:556). But this cannot be right either. As a logicist, Frege *cannot* hold that logic is indifferent to particular differences between objects: if arithmetic is to be reducible to logic, then logic must be capable of distinguishing the number 7 from the number 6.<sup>26</sup>

I suggest that for Frege, the generality of logic amounts to its general *applicability* to thought as such, whatever its topic.

**(Gen-6)** To say that logic is general or topic-neutral is to say that it provides norms for thought *as such*.

That is, Frege’s “generality” is 1-formality. I will argue in chapter 4 that Kant conceives of the generality of logic in this way, too.

In sum, there are three coherent ways of construing the generality or topic-neutrality of logic: as 1-formality, as 2-formality, and as 3-formality. The upshot is that everything we learn in this investigation of the “formality” of logic can be applied to discussions of logic’s generality or topic-neutrality as well.

<sup>26</sup>For more argument that Frege’s logic is not 2-formal, see section 5.3, below.

### 3.6 Philosophical significance of the three notions

Let us return briefly to the three philosophical *uses* for a demarcation of logic we considered in section 1.2. Here I will focus on logicism, with the aim of showing how it *matters* in what sense (if any) logic is formal.

If logic is 1-formal, then a proof that mathematics reduces to logic would show that mathematics is normative for thought as such, and hence independent of human sensibility. This is how Frege and Russell think of their logicist projects. For Russell, the point of logicism is to refute what he sees as the dangerous idealist doctrine that mathematical knowledge is conditioned by subjective facts about human sensibility, and hence only *conditionally* (not absolutely) true (see Hylton 1990a). For Frege, the point is to clarify the epistemological basis for mathematical knowledge—in particular, to establish its independence from sensibility (cf. Weiner 1990:ch. 2).<sup>27</sup> Neither Frege nor Russell takes logicism to show that mathematics lacks content. Russell says that Kant “rightly perceived” that mathematical truths are synthetic (1903:457), and Frege, though he takes mathematics to be analytic, redefines “analytic” so that analytic propositions can extend our knowledge and have content (FA:§88, §3).

Logicism takes on a very different significance if logic is demarcated as 3-formal, as it is by the positivists. In that case, to say that mathematics reduces to logic is to say that mathematics *lacks content* and says nothing about the real world.<sup>28</sup> As we have seen, Carnap and the positivists see in the 3-formality of logic and mathematics a way to account for a priori mathematical knowledge in an empiricist framework. Mathematics can be known *a priori*, on their view, because it represents no constraint on reality; it is not real knowledge at all.<sup>29</sup>

Finally, if logic is demarcated as 2-formal, then a Platonist logicism such as Frege’s—in

---

<sup>27</sup>Demopoulos 1994 and Tappenden 1995 argue that Frege’s main concern in giving logical proofs of arithmetical propositions is not to secure them from doubt, but to establish their general applicability (even outside the realm of the intuitable).

<sup>28</sup>For an especially vigorous statement of this view, see Ayer 1946:ch. 4.

<sup>29</sup>For the contrast between Russellian and positivistic logicism, see Hylton 1990a:144-5.

which numerals are taken as names for objects—becomes *impossible*, by definition. On the other hand, a non-Platonist type of logicism becomes *trivial*, since virtually any mathematical predicate or functor can be defined as a permutation-invariant logical constant (Sher 1991:132; cf. Tarski 1966:152-3, section 3.2, above). One of the accomplishments of Frege’s logicism is the logical definition of *finiteness*; but on the permutation invariance approach, “there are finitely many...” is already a logical constant. Sher 1991 embraces this consequence of her view:

The bounds of logic, on my view, are the bounds of mathematical reasoning. Any higher-order mathematical predicate or relation can function as a logical term, provided it is introduced in the right way into the syntactic-semantic apparatus of first-order logic. (xii-xiii)

But this consequence ought to make us wonder whether the notion of logicality Sher is explicating has anything to do with the notion of logicality presupposed by the original logicists. She claims that both the old logicism and the kind of logicism that would follow trivially from her account of logic “are based on the equation that being mathematical = being formal = being logical” (133). But it is not clear that anything more than the *word* “formal” connects them.

It matters a great deal, then, *in what sense* (if any) logic is formal.

### 3.7 A puzzle

I have proposed that talk of the “formality” of logic (and equally of its “generality” or “topic-neutrality”) moves between three quite different notions. To some extent, all three notions influence our “intuitions” about logicality: the intuitions to which philosophers appeal in motivating their demarcation proposals. In arguing that second-order logic cannot be logic because of its substantive mathematical content, one supposes that logic must be 3-formal; in arguing that arithmetic cannot be logic because of its special objects, one supposes that logic must be 2-formal; in arguing that quantum logic cannot be logic because it does not provide norms for thought as such, but only for thought in a quantum world, one supposes

that logic must be 1-formal. Distinguishing these three notions, I hope, will bring some light to the intractable debates about the bounds of logic.

But not just distinguishing them: it is also useful to ask how the three notions are related. And here there is a puzzle. Why are there three distinct notions of logical formality? How did they come to be confused under a single name? What are the connections between them? And how did they all come to be connected with logicity? Or are there several distinct notions of logicity in the tradition, corresponding to the different notions of formality? The only way to answer these questions is through conceptual archaeology. In chapter 4, I will argue that we need to start digging with Kant.

## Chapter 4

# KANT AND THE FORMALITY OF LOGIC

It is in Kant's critical philosophy, in which the distinction of matter and form is thoroughly grasped, that *formal* logic is first sharply separated out; and properly speaking, it stands and falls with Kant. However, many who otherwise abandon Kant have, at least on the whole, retained formal logic. (Trendelenburg 1870:15)

In this chapter, I will argue that contemporary logical hylomorphism traces back to Kant and cannot be understood apart from its Kantian pedigree. This point needs making, because the doctrine that logic is distinctively formal has become so thoroughly familiar that we no longer think of it as a characteristically Kantian doctrine, like transcendental idealism or the categorical imperative. Instead, it is often considered a “traditional” view, of which Kant is merely a particularly clear expositor.<sup>1</sup>

But in fact, Kant seems to have been the first modern philosopher to demarcate logic (or a significant portion thereof) by its formality.<sup>2</sup> Kant does not take his logical hylomorphism

---

<sup>1</sup>For example: “*As traditionally conceived*, logic is concerned with the form rather than the content of judgment. A sharp distinction between the form and content of judgment was to explicate the sense in which logic abstracted from the content of the claims of the special sciences. So, *for example*, Kant remarks: ‘That logic should have been thus successful is an advantage which it owes entirely to its limitations, whereby it is justified in abstracting. . . from all objects of knowledge and their differences, leaving the understanding nothing to deal with save itself and its form.’” (Ricketts 1986:81, emphasis added)

<sup>2</sup>This point was more widely acknowledged in the nineteenth century: see Trendelenburg 1870:15

from any of his modern predecessors: he self-consciously adopts it, *against* the current of his time, for his own purposes. Indeed, as I will show, he does not espouse it until the beginning of his critical period.

This, I will argue, is no accident: Kant's conception of logic is essential to his transcendental idealism; it is motivated by and articulated within the framework of that doctrine. This is not to say that it cannot be maintained in a different context. Perhaps it can be, but in that case it must be given a fresh motivation and defense. If we do not share Kant's reasons for thinking of logic as distinctively formal, we must give our own reasons, or reconsider the view. Unless we are willing to defend transcendental idealism, I suggest, we should be wary of logical hylomorphism.

The argument will proceed as follows. In section 4.1, I will look at Kant's characterization of (pure general) logic as "general" and "formal." By "general," I will argue, Kant means *1-formal*, while by "formal," he means *3-formal*. I will then claim that formality is not, for Kant, a defining feature of logic, but rather a substantive consequence of the generality of logic, given Kant's other philosophical commitments.

I give three different kinds of argument for this claim: exegetical, historical, and philosophical. The exegetical argument (section 4.1.4) exhibits passages in which Kant *argues* for the formality of logic on the basis of its generality. If formality were definitory of logic, argument would be out of place here.

The historical argument has two parts. In the first (section 4.2), I argue that Kant's claim that logic is formal (i.e., 3-formal) is original: he does not take it from any of his predecessors. If Kant had defined logic (in part) by its formality, he could justly have been charged with "changing the subject" in his disputes with the neo-Leibnizians over whether logic and conceptual analysis can yield knowledge of reality. It is because he defines logic in a way consistent with the earlier tradition that he can disagree with them about its formality.

---

(quoted above), De Morgan 1858:76 (quoted on page 131, below), and Mansel 1851:ii, iv (quoted on page 132, below). In the twentieth century I have found it only in Scholz 1931:2 and Sluga 1980:12, who both note that "formal logic" is a Kantian coinage.

In the second part of the historical argument (section 4.3), I show that Kant himself did not always characterize logic as distinctively formal. In his earlier logic lectures and marginal notes, we find little of the logical hylomorphism that pervades his later work, and some remarks that are flatly incompatible with it. The evidence suggests that Kant’s logical hylomorphism emerged during the years 1773-5—the very years in which Kant arrived at the central tenets of his critical philosophy.

Thus the historical evidence suggests a connection between Kant’s logical hylomorphism and his transcendental idealism. In the philosophical argument (section 4.4), I try to say what this connection is. I provide an argument, from Kantian premises, for the claim that if logic is to be general (1-formal), it must also be formal (3-formal). The argument shows how Kant’s philosophy of logic is entangled with his more general philosophical views.

Finally, in section 4.5, I look briefly at the *influence* of Kant’s philosophy of logic on the later tradition. I show how Kant’s claim that logic is formal gets transformed from a substantive thesis about logic into part of the definition of logic, so that even philosophers who do not share Kant’s philosophical views must find a sense for it. It is in this way, I suggest, that the word “formal” comes to be separated from 1-formality and attached to 2-formality and 3-formality as well.

## 4.1 Kant’s characterization of logic

### 4.1.1 Kant’s taxonomy of logics

First, a caveat. In the first *Critique*, it is only *pure general logic*, not logic *simpliciter*, that Kant takes to be formal. Logic, in the broadest sense of the term, is “the science of the rules of the understanding in general,” as opposed to aesthetic or “the rules of sensibility” (KrV:A52/B76). It has two main branches: *general logic*, which concerns the “absolutely necessary rules of thought without which there can be no employment whatsoever of the understanding,” and *special logics*, which concern “the rules of correct thinking as regards a certain kind of objects,” that is, the objects of a particular science (A52/B76). This dis-

inction corresponds to the *Jäsche Logic*'s distinction between *necessary rules* of cognition, "without which no use of the understanding would be possible at all," and *contingent rules* of cognition, "without which a certain determinate use of the understanding would not occur," and "which depend upon a determinate object of cognition" (JL:12).<sup>3</sup> Contingent rules are contingent in the sense that they are not universally applicable to thought *as such*:

Thus there is, for example, a use of the understanding in mathematics, in metaphysics, morals, etc. The rules of this particular, determinate use of the understanding in the sciences mentioned are contingent, because it is contingent whether I think of this or that object, to which these particular rules relate. (JL:12)<sup>4</sup>

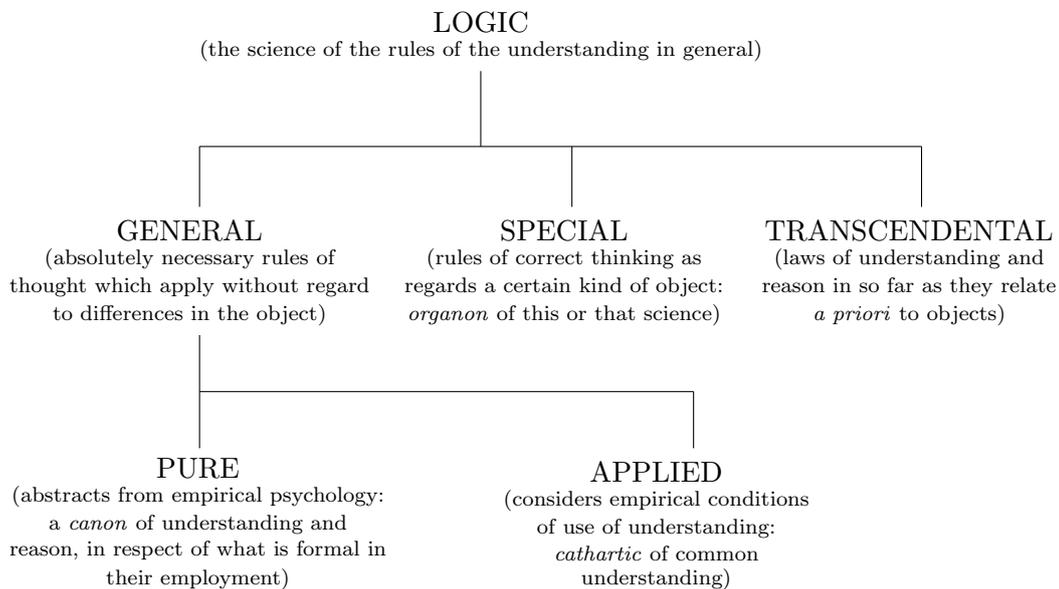
General logic, in turn, is divided into *pure* and *applied*: pure general logic is an *a priori* science that abstracts entirely from contingent features of human psychology (memory, habits, etc.), while applied general logic considers empirical psychological principles that affect actual human thought. General logic is also distinguished from *transcendental logic*, which does not abstract *entirely* from objects and the content of thought, but "concerns itself with the laws of understanding and of reason solely in so far as they relate *a priori* to objects" (A57/B82). Kant does not say whether he takes transcendental logic to be a *special* logic, but there is good ground for thinking he does. He seems to regard the restriction of transcendental logic to objects capable of being given in human sensibility as a *domain restriction*, like the restriction of geometry to spatial objects. Thus, for instance, he characterizes the use of the understanding beyond the bounds set by transcendental logic as directed toward "... objects *without distinction*—upon objects which are not given to us, nay, perhaps cannot in any way be given," that is, toward "objects in general" (A63/B88, my emphasis). Similarly, in the *Jäsche Logic*, he says that transcendental logic represents the object "as an object of the mere understanding," while general logic "deals with all objects in general" (JL:15). And in R:1628 (at 44.1-8), Kant uses "objects of experience"

<sup>3</sup>See section 3.1, above.

<sup>4</sup>The JL, which was put together from Kant's notes and annotations, must be used with care as a source for Kant's views on logic. See Longuenesse 1998:81 n. 1 and Young's introduction to Kant 1992:xviii. In general, I take it to be a fairly reliable guide. Much of its content can also be found in other transcripts of Kant's logic lectures, or in Kant's published works.

as an example of a particular domain of objects that would require special rules (presumably, those of transcendental logic)—as opposed to the “rules of thinking *überhaupt*” contained in general logic.<sup>5</sup> All of these passages suggest that Kant takes transcendental logic to be a special logic. However, the evidence is not conclusive. Kant characterizes the special logics as *organa* of the various sciences (A52/B76), and in places, he denies that transcendental logic is an *organon*.<sup>6</sup> For now, we will leave the question open (see figure 4.1).

Figure 4.1: Kant’s taxonomy of logics.



Of these divisions, only pure general logic corresponds to what we now call “logic.” The special logics contain the principles of particular sciences, insofar as these can be seen as defining “forms of thought” characteristic of these sciences (JL:18, PL:508). Applied logic,

<sup>5</sup>“Aber nachdem die objecte Verschieden seyn, müssen auch verschiedene Regeln des Denkens seyn, z. B. Andere Regeln im Gegenstand der Erfahrung als im Gegenstand der blossen Vernunft (Tugend), andere Regeln des Verstandes vor äussere Erfahrung als vor innere. Jede Wissenschaft hat ihre besondere Regeln. // Es muss aber doch auch eine Geben, die vor allen Wissenschaften vorhergeht und die Regeln des Denkens überhaupt enthält. Hier muss von allem unterschiede der objecte abstrahirt werden.”

<sup>6</sup>For example, at R:2162: “In der transcendentalen Logik ist die materie allgemein bestimmt und unterschieden; daher criterion der warheit, aber kein organon.” But compare A1011/B245, where the Critique of Pure Reason is said to be “a special science” aiming at an “organon of pure reason.”

on the other hand, is properly a part of psychology. In the *Jäsche Logic*, Kant denies that it ought to be called logic at all, on the grounds that “[i]t is a psychology in which we consider how things customarily go on in our thought, not how they ought to go on” (JL:18). Only pure general logic and transcendental logic have the general applicability and normative character we associate with logic. But transcendental logic concerns itself with the relation between thought and its objects, and hence with intuition and sensibility: we would call it epistemology or semantics (in a broad sense), not logic. Even Kant often uses the term “logic” in the stricter sense of “pure general logic” (e.g., KrV:B ix, A61/B86, A598/B626; JL:13, 14). And when logic (rather than critical philosophy) is his topic, he *defines* logic in a way that excludes transcendental logic.<sup>7</sup> In what follows, I will use “logic” to mean “pure general logic,” unless I explicitly specify otherwise.

#### 4.1.2 What does Kant mean by “general”?

What does Kant mean by the “general” in “general logic”? General logic, Kant says, “. . . treats of understanding without any regard to difference in the objects to which the understanding may be directed” (A52/B76; cf. A54/B78; JL:12, 50-51; GMS:387). This suggests that by “general” he means “indifferent to the particular identities of objects,” i.e., *2-formal*.<sup>8</sup>

However, this kind of generality does not distinguish logic from arithmetic and algebra. Kant holds that arithmetic and algebra, too, “. . . [abstract] completely from the properties of the object (*Beschaffenheit des Gegenstandes*) that is to be thought in terms of. . . a concept of magnitude” (A717/B745).<sup>9</sup> An arithmetical truth like “ $5+7=12$ ” is applicable to cows,

<sup>7</sup>In the JL, having just said that transcendental logic considers the object “as an object of the mere understanding,” whereas general logic “. . . deals with all objects in general,” Kant defines *logic* as “. . . a science a priori of the necessary laws of thought, not in regard to particular objects, however, but to *all objects in general* . . .” (JL:16, my emphasis).

<sup>8</sup>For the different notions of formality, see chapter 3, above. For the connection between these notions and the generality of logic, see section 3.5.

<sup>9</sup>Friedman 1992:108 argues convincingly that although only algebra is explicitly mentioned in this passage, the point applies to arithmetic as well, since both “[construct] magnitude as such (*quantitas*)” (A717/B745). In saying that algebra abstracts from the constitution (*Beschaffenheit*) of the object, Kant is not alluding to the fact that it operates with variables instead of determinate

murders, inches, and triangles alike: it considers objects as *units*, abstracting entirely from their distinguishing characteristics. Kant does not think of the numerals “5,” “7,” “12” as names of objects. On his view, mathematics is about *forms of* empirical objects, not about “mathematical objects.”<sup>10</sup> But while geometry is limited in its application to a special domain of such objects—the objects capable of being given in space—arithmetic and algebra have no corresponding limitation (Friedman 1992:113). As “techniques of calculation,” they are “independent of the specific nature of the objects whose magnitudes are to be calculated” (113). Pure algebra and arithmetic do not assume that the magnitudes to which they apply are spatial, or even temporal (114-5).<sup>11</sup> On Friedman’s view, the role of pure intuition in arithmetic and algebra is not to provide these sciences with objects, but to give content to the *concept* of magnitude—that is, of successive iteration of units—on which they are based (122). Time is “a universal source of *representations* for the number series,” not (as on Parsons’ view) “a universal source of *models* for the numbers” (122 n. 46, emphasis added; cf. Parsons 1969:140). The concept of magnitude must get its content from the iteration of operations in time, because the successive iteration of units cannot even be *represented* using the logic available to Kant. We post-Fregeans can represent the indefinite extendibility of the number series by saying “for every number, there is a successor.” Kant, quantities, but to the fact that it concerns itself with magnitude as such (*quantitas*), rather than particular magnitudes (*quanta*), like the spatial magnitudes considered in geometry. In this respect arithmetic and algebra are alike. Friedman also suggests, more contentiously, that algebra for Kant is distinguished from arithmetic primarily by its concern with irrational magnitudes (and not, in the first instance, by its use of variables) (108-112).

<sup>10</sup>See KrV:A239-40/B298-9, B147, A224/B271, Thompson 1972-3:338-342, and Friedman 1992:101.

<sup>11</sup>“The application of the science of *quantity*, unlike that of the science of geometry, is therefore not limited to the specific—that is, spatial—character of our intuition; in this sense it provides us with the concept of a thing in general” (114). In a letter to Schultz, Kant calls *quantity* “a concept of a thing in general by determination of magnitude” (quoted in Parsons 1969:134). In view of KVR:A719/B747, Parsons connects this phrase with Kant’s characterization of the categories as “concepts of an object in general” (A51/B75, A93/B126, B128) and with the “intellectual synthesis” of B151 (134-5; cf. Thompson 1972:338). “Thing in general” is used in the Amphiboly of Concepts of Reflection for a thing thought in abstraction from the conditions of intuition (A279/B335, A283/B339). Friedman suggests that it might be more correct to speak of “the concept of *an object of intuition in general*” (114 n. 34), since magnitude has to do with “the synthesis of the homogeneous in an intuition in general...” (B162). Even with this qualification, however, the concept of *quantity* has not been limited to specifically spatiotemporal objects. See also Pr:§§39, 45 and KrV:A53/B51.

lacking the apparatus of iterated quantifiers, is forced to appeal to the intuitive idea of iteration in *time* to represent the same idea:

The concept of magnitude in general can never be explained except by saying that it is the determination of a thing whereby we are enabled to think how many times a unit is posited in it. But this how-many-times is based on successive repetition, and therefore on time and the synthesis of the homogeneous in time. (KrV:A241/B300)

If Friedman is right, then logic cannot really be distinguished from algebra and arithmetic on the grounds of its 2-formality. Algebra and arithmetic are 2-formal, too: they abstract just as surely as logic from particular differences between objects. (This is not to say that they yield any *knowledge* of objects beyond sensory experience. But neither does logic. Such knowledge is not possible at all, on Kant's view.)

These considerations tell decisively against construing Kant's "generality" as 2-formality. For in characterizing logic as general, Kant clearly intends to be *distinguishing* it from mathematics. In the *Jäsche Logic*, for instance, Kant defines general logic as the science of the *necessary* rules of the understanding, "without which no use of the understanding would be possible at all," contrasting these with various *contingent* rules of the understanding, "without which a certain determinate use of the understanding would not occur" (JL:12). One of his examples of such a "particular, determinate use of the understanding" is the "use of the understanding in mathematics" (ibid.). Arithmetic and algebra, then, are not "general" in the sense that logic is. Though they may not have their own special objects, and though their concepts may not distinguish between the particular identities of objects, their laws are applicable only to a "particular, determinate use of the understanding," not to all uses of the understanding. Because the content of the concept of magnitude depends on sensibility, while thought is intelligible apart from sensibility, the norms governing this concept (i.e., the laws of arithmetic and algebra) cannot be norms for thought *as such*. They cannot be laws of "...the general ...employment of the understanding" (A52/B76).

It appears, then, that when Kant speaks of the "generality" of logic, he means its *1-formality* (constitutive normativity for thought *as such*), not its 2-formality. This is borne

out by his characterization of general logic as the science of “...the absolutely necessary rules of thought without which there can be no employment whatsoever of the understanding,” as opposed to “...the rules of correct thinking as regards a certain kind of objects” (KrV:A52/B76; cf. JL:13). Kant is explicit that he does not mean rules by which the understanding *does* proceed, but rules governing how it *ought* to proceed (JL:14). The point is that no activity that is not held accountable to these rules can *count* as thought, and not that there cannot *be* thought that does not conform to these rules.<sup>12</sup>

### 4.1.3 What does Kant mean by “formal”?

Kant invokes *formality* both in distinguishing general logic from special logics (A54/B78) and in distinguishing general logic from transcendental logic (A556/B80). Pure general logic, Kant holds, “...abstracts from all content of knowledge, that is, from all relation of knowledge to the object, and considers only the logical form in the relation of any knowledge to other knowledge; that is, it treats of the form of thought in general” (A55/B79). This claim is repeated in many places throughout Kant’s works, with minor variations (no doubt all intended as equivalent). General logic abstracts entirely from the objects themselves (KrV:B ix; JL:51, 16)—or more properly, from relation of thought to the objects (KrV:A55/B79), that is, from the content of thought (KrV:A54/B78, A55/B79, A70/B95; JL:94).<sup>13</sup> Thus it is the science of “the mere form of thought” (JL:13) and deals “...with nothing but the mere form of thought” (KrV:A54/B78; cf. A55/B79, A56/B80, A70/B95, A131/B170). Its laws are “*without content* and merely formal” (A152/B191, emphasis added).

I suggest that the sense of formality at issue here is *3-formality*: abstraction from

<sup>12</sup>Here I concur with Smit 1999:214. Putnam 1994 claims to find in Kant the idea that “illogical thought is not, properly speaking, thought at all” (246; cf. Conant 1991). I see little basis for this reading of Kant, but there are a few passages that might support it. At KrV:A150/B189, Kant says that self-contradictory judgments are “nothing” (*nichts*: Kemp-Smith softens this to “null and void”). (I owe this observation to Steve Engstrom.) And in the *Vienna Logic*, Kant is reported to have said: “Logic has the peculiarity that the subjective laws are also objective rules, because the universal rules are the sole condition of our thought” (VL:791).

<sup>13</sup>For the equation between content and relation to objects, see KrV:A58-9/B83, A63/B87.

all semantic content of the representations used in thought.<sup>14</sup> What this means can be clarified by the consequences Kant draws from it. Because logic is formal, he says, it affords no knowledge of objects:<sup>15</sup>

... if I separate understanding from sensibility to obtain a pure understanding, then nothing remains but the mere form of thinking without intuition, by which form alone I can know nothing definite and consequently no object. (Pr:§57)

For logic teaches us nothing whatsoever regarding the content of knowledge, but lays down only the formal conditions of agreement with the understanding; and since these conditions can tell us nothing at all as to the objects concerned, any attempt to use this logic as an instrument (*organon*) that professes to extend and enlarge our knowledge can end in nothing but mere talk... (KrV:A61/B86, cf. A60/B85)

That is, logic (properly understood) does not claim to tell us how things are with objects in the world. It delivers no knowledge of fact—not even of the most abstract and general facts (facts about identity or existence, for example). In this it contrasts with the *a priori* mathematical sciences, which purport to give us real knowledge about empirical objects (though only as to their sensible *forms*). Whereas a transcendental deduction is needed to ensure the applicability of mathematics to empirical objects, no comparable assurance is needed for logic, because logic (unlike mathematics) makes no claims about objects. The point is not just that it does not tell us about *actual* objects: it does not tell us about *possible* objects, either.<sup>16</sup> The “logical possibility”—that is, freedom from self-contradiction—of a concept does not suffice for its objective validity or “real possibility”—that is, its relation to some definite object “in the sum of all possibilities” (KrV:B xxvi n.). For example, the concept *two-sided figure* is free from logical contradiction, yet no possible object is a two-sided figure.

So much for the negative aspect of Kantian “formality.” What about the positive aspect?

<sup>14</sup>It is certainly not 2-formality, since algebra and arithmetic are 2-formal but not “merely formal” in Kant’s sense. Nor is it 1-formality, or “generality.” As we will see in section 4.1.4, below, Kant regards “generality” and “formality” as distinct, though intimately related properties.

<sup>15</sup>The qualification “of objects” is essential, because Kant allows that logical laws afford knowledge of the truth of *analytic* propositions (A151/B190), which make explicit the contents of *concepts*.

<sup>16</sup>In his reply to Eberhard, Kant says that logical principles “... completely abstract from every-thing concerning the possibility of the object...” (E:194).

What does it mean to say that logic “treats of the form of thought in general”?

One should not picture “the form of thought in general” as a kind of mental glue by means of which representations are stuck together, and logic as a quasi-psychological investigation of its adhesive properties. The form of thought is not any kind of thing (not even a mental thing). It is, rather, a set of *norms*: in fact, the laws of logic themselves. If this is right, then Kant’s “positive” characterization of formality adds nothing substantive to the “negative” characterization. To say that logic treats of the form of thought in general is to say that it treats of the laws of logic. Let me explain why.

Kant characterizes the understanding as “the faculty of rules” (A13/B171) or “the source of rules” (JL:12). The reason is that the work of the understanding—the “spontaneous” production of concepts, judgments, and inferences—consists in the institution of norms. A concept, for Kant, is not an image or a collection of habitually associated images. Instead, it is constituted by a *rule* which determines how it can be *correctly* applied in particular cases. To say that the concept *animal* is one of the marks of the concept *cat*, for example, is not to say that when we think of a cat, we think of an animal; it is to say that applying the concept *cat* to a particular object of intuition *commits* one to applying the concept *animal*. Thus

... a concept is always, as regards its form, something universal which serves as a rule. ... The concept of body, in the perception of something outside us, necessitates the representation of extension, and therewith representations of impenetrability, shape, etc. (KrV:A106)<sup>17</sup>

In a similar way, a judgment is a rule for the relation of representations. If I judge that all cats are cunning, I institute a *rule* by the lights of which I am *required* to apply the concept *cunning* to anything to which I apply the concept *cat*:

Judgments, when considered merely as the condition of the union of given representations in a consciousness, are rules. (Pr:§23).<sup>18</sup>

<sup>17</sup>As Longuenesse explains, a concept “... is a rule in that thinking an object under a concept provides a reason to predicate of this object the marks that define the concept” (1998:50). A concept is also a rule for the unification of the manifold of intuition. See Longuenesse 1998:48-52 for a much more nuanced discussion of the rulishness of concepts.

<sup>18</sup>See Longuenesse 1998:93-5 for a fuller discussion of judgments as rules. Note that the rules that

The essential rulishness of judgments (and derivatively of the concepts whose contents they explicate) manifests itself most clearly when they are used as major premises of syllogisms: “the syllogism is itself nothing but a judgment made by means of the subsumption of its condition under a universal rule (the major premiss)” (KrV:A307/B364).

Broadly speaking, Kant thinks of this normative aspect of concepts and judgments—their rulishness—as their *form*. Thus, the form of concepts is *universality* (JL:§2, 91), while the form of judgments is the determination of the *relation* between representations: that is, “. . . the determination of the way that the various representations [i.e., the matter of the judgment] belong, as such, to one consciousness” (JL:§18, 101). In a categorical judgment, for example, the subject and predicate are the matter, while the normative relation between them (e.g., “one ought to apply B to everything to which one applies A”) is the form (JL:105). The logical functions of judgment (as displayed in the table of judgment) are the different possible normative relations: “the various modes of uniting representations in consciousness” (Pr:§22).

Every concept and judgment has a form, then—its “rulishness.” But what is “the form of thought in general”? The form of thought in general is that which makes thought possible: the “. . . formal conditions of all judgments in general (and hence of all rules in general)” (Pr:§23). But what are the conditions of all rules in general? What must be in place before the understanding can institute a norm—say, a the norm that one ought not apply A and B to the same thing? One needs a way of *indicating* incompatibility—say, the symbol “ $\perp$ ”—but that is not fundamental. For we must then ask in virtue of what the symbol “ $\perp$ ” indicates incompatibility. And the answer is clear: a symbol “ $\perp$ ” can only indicate incompatibility if there is a particular kind of norm for its use: a norm like

( $\perp$ -rule) If  $A \perp B$ , then one ought not apply A and B to the same thing.

Hence the condition for the understanding’s activity is a set of norms that make it possible to institute norms of incompatibility, universality, and so on. These norms—which we can judgments are can be either “objective” and necessary for every consciousness or “subjective” and necessary for a particular consciousness at a particular time (Pr:§22-3). General logic abstracts from this distinction.

now recognize as the laws of logic—are the rules according to which the source of rules itself proceeds (JL:12-13). Thus the form of thought in general, of which logic treats, is nothing other than the necessary rules for the employment of the understanding (JL:13), or norms for thought *as such*. It is in this light that we should understand Kant’s claim at JL:13 that the “science of the necessary laws of the understanding and of reason in general” and the science of “the mere form of thought as such” are “one and the same.”

The positive part of Kant’s claim that general logic “. . . abstracts from all content of knowledge . . . and . . . treats of the form of thought in general,” then, adds nothing to his characterization of it as “general” (that is, as 1-formal). It is the negative part of the claim that adds something new: namely, that the norms for thought *as such* are in no sense about the world, abstract entirely from the content of thought, and can give us no knowledge of objects.

#### 4.1.4 Generality, formality, and Kant’s demarcation of logic

As we have seen, Kant characterizes logic as both “general” and “formal.” We can understand the difference between these two characterizations in terms of our three notions of logical formality. In characterizing logic as “general,” Kant is characterizing it as 1-formal, or normative for thought as such. In characterizing logic as “formal,” he is characterizing it as 3-formal—as abstracting entirely from semantic content (or, as he often puts it, “relation to objects”). But what is the status of these two characterizations of logic? Does Kant regard generality and formality as independent defining features of logic? Or is one primary, the other derivative?

In what follows, I will argue that Kant takes logic to be demarcated by its *generality*, and that formality is not an independent defining feature of logic, but a substantive consequence of its generality, given certain other Kantian assumptions. That is, although Kant thinks that logic must be formal, he does not take formality to be part of the *essence* of logic. That a general logic must also be formal is a *substantive claim* that requires justification, not a matter of semantics. I call this claim

**Kant's Thesis:** General (i.e., 1-formal) logic must also be formal (i.e., 3-formal). That is, "...the universal and necessary rules of thought in general can concern merely its *form* and not in any way its *matter*" (JL:12).

My argument that formality is not part of Kant's *concept* of logic, but a substantive consequence of logic's generality, has three main parts. The first is purely exegetical. If formality were part of the *concept* of (pure general) logic, we would expect it to play a role in Kant's taxonomy of "logics." But it does not. In the first *Critique*, general logic is distinguished from special logics by its *generality*: it

...contains the absolutely necessary rules of thought without which there can be no employment whatsoever of the understanding. It therefore treats of understanding without any regard to difference in the objects to which the understanding may be directed. (A52/B76)

*Pure* general logic is distinguished from applied general logic by its abstraction from the empirical conditions of its use:

In the former we abstract from all empirical conditions under which our understanding is exercised, i.e. from the influence of the senses, the play of imagination, the laws of memory, the force of habit, inclination, etc., and so from all sources of prejudice, indeed from all causes from which this or that knowledge may arise or seem to arise. For they concern the understanding only in so far as it is being employed under certain circumstances, and to become acquainted with these circumstances experience is required. (A53/B77)

The distinction between pure and applied echoes the distinction between general and special: just as special logics concern themselves not with thought as such but with thought about a particular *objective* domain, so applied logics concern themselves not with thought as such but with thought under particular *subjective* conditions of employment. Neither of these distinctions is a distinction between formal and material.

It is true that, immediately after making these distinctions, Kant invokes formality in describing the differences between pure and applied logics and between general and special logics:

Pure general logic has to do, therefore, only with principles *a priori*, and is a *canon of understanding* and of reason, but only in respect of what is formal

in their employment, be the content what it may, empirical or transcendental. (A53/B77)

... There are therefore two rules which logicians must always bear in mind, in dealing with pure general logic:

1. As general logic, it abstracts from all content of the knowledge of understanding and from all differences in its objects, and deals with nothing but the mere form of thought. (A54/B78)

But these passages are best read as displaying consequences of the taxonomy Kant has just provided, not as giving further *grounds* for the distinctions between pure and applied, general and special. (Note the “therefore”s in both passages.) Pure general logic must be formal, not by definition, but because otherwise it cannot be pure and general.

Kant makes much of the formality of general logic in distinguishing it from *transcendental logic*:

General logic, as we have shown, abstracts from all content of knowledge, that is, from all relation of knowledge to the object, and considers only the logical form in the relation of any knowledge to other knowledge; that is, it treats of the form of thought in general. But since, as the Transcendental Aesthetic has shown, there are pure as well as empirical intuitions, a distinction might likewise be drawn between pure and empirical thought of objects. In that case we should have a logic in which we do not abstract from the entire content of knowledge.

It might be thought that, since Kant invokes formality to distinguish transcendental logic from general logic, formality must be part of the *concept* of general logic. This reasoning would be cogent if transcendental logic were, like general logic, *general* (i.e., 1-formal). But it is not. Its norms are “rules of the pure thought of an object”, not norms for thought as such. Not all thought is, for Kant, “pure thought of an object”: ordinary empirical thought is not pure, and some thought (e.g., in morals and religion) is not “of an object” at all. Transcendental logic is therefore a *special* logic (see section 4.1.1, above). There is no need, then, to add anything to the definition of general logic to separate it from transcendental logic; generality suffices.<sup>19</sup> But since formality is a consequence of generality,

<sup>19</sup>Indeed, the way Kant begins the paragraph—“General logic, *as we have shown*, abstracts from all content of knowledge ...”—would be very odd if the connection between formality and general logic were definitional.

it is appropriate for Kant to invoke it here in explaining the difference between formal and transcendental logic.

Kant's demarcation of general logic in the *Jäsche Logic* also suggests that the formality of logic is a substantive consequence of its generality, not an independent defining characteristic:

[1] If now we put aside all cognition that we have to borrow from *objects* and merely reflect on the use just of the understanding, we discover those of its rules which are necessary without qualification, for every purpose and without regard to any particular objects of thought, because without them we would not think at all. [2] Thus we can have insight into these rules *a priori*, *i.e.*, *independent of all experience*, because they contain merely the conditions for the use of the understanding in general, *without distinction among its objects*, be that use *pure* or *empirical*. [3] **And from this it follows** at the same time that the universal and necessary rules of thought in general can concern merely its *form* and not in any way its *matter*. [4] Accordingly, the science that contains these universal and necessary rules is merely a science of the form of our cognition through the understanding, or of thought. (JL:12, boldface emphasis added)

In [1], Kant is referring to the *generality* of logical laws: that is, their normativity for thought as such. In [2] and [3], he draws two further conclusions from the generality of logical laws: they must be knowable *a priori* and they must be purely *formal*.<sup>20</sup> [4] sums up: a general logic must also be formal.

All of this exegesis gives *prima facie* plausibility to my claim that Kant's commitment to the formality of logic rests on a substantive philosophical thesis (Kant's Thesis), not on semantics. But the strongest arguments are yet to come. First I will show that it wouldn't have made *sense* for Kant to build (3-)formality into the definition of logic, given

<sup>20</sup>The text—"Und hieraus folgt zugleich"—might also be construed as saying that the formality of logic follows from its *a priori* knowability. But the interpretation I have suggested seems more natural (especially in view of "zugleich") and is certainly preferable philosophically. It does not follow from the *a priori* knowability of a law that it concerns merely the form of thought "and not in any way its *matter*": otherwise, general logic would be the only *a priori* science. My interpretation is confirmed by one of the Reflexionen (R:1620, at 40.23-5), where Kant writes: "Eine allgemeine Verstandeslehre trägt also nur die nothwendige Regeln des Denkens vor ohne unterschied der obiecte, d. i. der Materie, worüber gedacht wird, *also* nur die Form des Denkens überhaupt und die Regeln, ohne welche gar nicht gedacht werden kan" (emphasis added). Here there is nothing about *a priori* knowability, but otherwise the point is the same; the inference is clearly from the generality of logic to its formality. For a similar inference, see R:2162: "If one speaks of cognition *überhaupt*, then one can be talking of nothing beyond the form."

his historical and dialectical context. For none of his predecessors characterize logic as 3-formal. Some, indeed, have views that require logic *not* to be 3-formal—views Kant is intent to refute. In building formality into the definition of logic, Kant would have been begging huge questions against his rationalist opponents, who held that logic and conceptual analysis can yield real knowledge about the world. Far better to read him as defining logic in a way that these opponents could accept, as the science of the norms generally binding on thought *as such*.

Second, I will show that Kant only begins to characterize logic as formal at around the time of his conversion to transcendental idealism. This suggests that he is driven to the claim that logic is formal by his broader philosophical commitments.

Finally, I will reconstruct an argument for Kant's Thesis, using Kantian premises. In this way, I will show how Kant's critical turn made it *necessary* for him to characterize logic as formal—not as a matter of definition, but as a substantive thesis of transcendental philosophy.

## 4.2 The originality of Kant's characterization

If the delineation of logic by its *formality* were traditional, not a Kantian invention, we should expect to find it in his modern predecessors. But in fact, I will argue, we do not: Kant's doctrine of the formality of logic is original. To be sure, one can find foreshadowings of the doctrine in Kant's predecessors. But not even transcendental idealism is spun out of whole cloth.<sup>21</sup> My claim is not that Kant's delineation of logic as formal is *wholly* novel (no idea is), but rather that it has as much claim to being a Kantian innovation as transcendental idealism or the categorical imperative. Let us survey the sources from which Kant might have picked up the idea.

---

<sup>21</sup>See de Vleeshauwer 1939.

### 4.2.1 The Wolffian school

A reasonable starting point is the official text for Kant's lectures on logic: Georg Friedrich Meier's *Auszug aus der Vernunftlehre*. Meier's text is a synopsis of the logic of the Wolffian school of neo-Leibnizian rationalists,<sup>22</sup> and like the other Wolffian philosophers, Meier does not characterize logic as distinctively formal. He defines logic as "a science that treats the rules of learned cognition and learned discourse" (§1), dividing this science in various ways, but never into a part whose concern is the *form* of thought. Although Meier follows tradition (e.g., Arnauld and Nicole 1662:218) in distinguishing between material and formal incorrectness in inferences (§360, cf. §§359, 395), he never makes the point that logic is concerned only with the formal variety. In fact, the distinction he draws between formal and material is simply skew to Kant's. In Meier's sense, material correctness amounts to nothing more than the truth of the premises, while formal correctness concerns the *connection* between premises and conclusion. But for Kant, to say that logic is formal is not to say simply that it is concerned with inferential connections (as opposed to the truth of premises), for some inferential connections will have material grounds.<sup>23</sup> The point is rather to restrict logic's concern to connections *that hold in virtue of the form of thought*, abstracting from content.

In characterizing logic as formal in his lectures, then, Kant is going beyond Meier's text.<sup>24</sup> This should not be surprising: Kant's characterization of pure general logic is developed in explicit opposition to the Wolffian conception of logic expounded by Meier. As Trendelenburg points out (1870:15), Wolff thinks that logic should be grounded in ontology and psychology, and that it is prior to them only in the order of learning.<sup>25</sup> Wolff argues

<sup>22</sup>In a capsule history of logic, Kant tells us that Meier's source was Baumgarten, who in turn concentrated the logic of Wolff, whose general logic is "the best we have" (JL:21). For the dominance of the Wolffian system in German universities of Kant's time, see de Vleeschauwer 1939:9-10.

<sup>23</sup>For example, there are many inferences in Euclidean geometry that are not licensed by Aristotelian logic. Kant takes these to be grounded in facts about the pure intuition of space and time (see Friedman 1992:chapter 1).

<sup>24</sup>The sceptical reader may examine all of Meier's uses of matter/form in relation to logic: §§128, 191, 309, 359, 360, 395, 399, 402, 428, 430, 514, 515, 516.

<sup>25</sup>"Methodus demonstrativa requirit, ut Logica post Ontologiam & Psychologiam tradatur (§. 90.); methodus autem studendi suadet, ut eadem omnibus philosophiae partibus ceteris praeponatur, con-

as follows:

For [logic] treats of rules, by which the intellect is directed in the cognition of every being (§61): the definition does not restrict it to a certain kind of being. Thus it ought to teach those [principles] we must attend to in cognition of things. But these should be derived from the cognition of being in general, which is taken from ontology (§73). It is plain, therefore, that principles should be sought from ontology for the demonstrations of the rules of logic.

Further, since logic expounds the method of directing the intellect in the cognition of truth (§61), it ought to teach the use of [the intellect's] operations in cognizing truth. But what the cognitive faculty is, and what are its operations, must be learned in addition from psychology (§58). Therefore it is also plain that principles should be sought from psychology for the demonstrations of the rules of logic. (Wolff L: *Discursus praeliminaris*, §89, my translation)<sup>26</sup>

For Kant, by contrast, (pure general) logic must abstract from both ontology and psychology:

1. As general logic, it abstracts from all content of the knowledge of understanding and from all differences in its objects, and deals with nothing but the mere form of thought.
2. As pure logic, it has nothing to do with empirical principles, and does not, as has sometimes been supposed, borrow anything from psychology. . . (KrV:A54/B78)

The contrast could not be clearer. Not only does Kant not take the delineation of logic as formal from the Wolffians, he invokes that doctrine to *distinguish* his conception of logic from theirs.

What is at issue between Kant and the Wolffians comes out very clearly in Kant's polemics against them. Kant takes Eberhard to task for conflating formal and material versions of the principle of sufficient reason. "Every proposition must have a reason" is, sequenter & Ontologiam atque Psychologiam praecedat (§. 88.)." (Wolff L: *Discursus praeliminaris*, §91)

<sup>26</sup> "Tradit enim ea regulas, quibus dirigitur intellectus in cognitione omnis entis (§. 61.), neque enim definitio eam ad certam entium speciem restringit. Docere igitur debet, quatenam sint ea, ad quae in rerum cognitione attendere tenemur. Haec autem ex generali entis cognitione derivanda, quae ex Ontologia hauritur (§. 73.). Patet itaque ad demonstrationes regularum Logicae petenda esse principia ex Ontologia. Porro cum Logica exponat modum dirigendi intellectum in cognitione veritatis (§. 61.); usum docere debet operationum ipsius in veritate cognoscenda. Quatenam vero sit facultas cognoscitiva, quatenam ipsius sint operationes, ex Psychologia addiscendum (§. 58.). Patet igitur porro, quod ad demonstrationes regularum Logicae petenda sint ex Psychologia principia."

Kant agrees, a formal logical principle (193), but “every *thing* has a reason” is not (194). He accuses Eberhard of trying to dupe the reader by using the ambiguous “*all* has a reason:”

He wishes to validate this concept of reason (and with it, unnoticed, the concept of causality) for all things in general, i.e., prove its objective reality without limiting it merely to objects of the senses, and thereby avoid the condition stipulated by the *Critique*, namely, the necessity of an intuition by means of which this reality is first demonstrable. (194-5)

Similarly, he accuses Baumgarten of taking the *logical* principle that “one cannot modify anything in that which pertains essentially to [a] concept, without at the same time giving up the concept itself” as a *metaphysical* principle extending our knowledge of objects: the principle that “*the essences of things are unalterable.*” “Philosophical novices” use this principle to

. . . reject, for example, the opinion of some mineralogists that silicon can be gradually transformed into aluminum. Only this metaphysical motto is a poor identical proposition which has nothing at all to do with the existence of things and their possible or impossible alterations. Rather it belongs entirely to logic. . . . (E:237 n.)

These passages are representative of Kant's general criticism of Wolffian metaphysics: it makes analytic, purely logical propositions *look like* real principles about objects by exploiting words that have (ambiguously) both a real use (applied to objects) and a merely logical use (applied to mere concepts) (E:237-8).

From the Wolffians' point of view, these criticisms are unfounded, since the distinctions Kant accuses them of conflating are spurious. The constitutive principles for thought as such are also the principles for *being* as such, for just the reason given in the passage from Wolff quoted above: since thought is about being, the norms for thought in general must depend on the most general truths about being. Logic, as the Wolffians conceive it, abstracts from *particularity*, but not from being and reality altogether. Allison 1973 describes Eberhard's position as follows:

In proceeding by way of abstraction, the understanding keeps its contact with the real. It only loses the features pertaining to individuals in their particu-

larity, what Eberhard calls the imageable or intuitive features, but it keeps the unimageable yet still conceivable features pertaining to all reality. (25)<sup>27</sup>

It would plainly do Kant no good, in this dialectical context, to *define* logic in a way that builds in formality. For, whether or not one uses the word “logic,” there remains a disputed question here, namely: must the science of norms for thought as such abstract *entirely* from content, or only from *particular* content? Kant plumps for the former answer, but his grounds for doing so are intimately entwined with his broader philosophical views (as we will see in section 4.4, below).

Thus Kant's doctrine that logic is formal is not merely absent from his Wolffian contemporaries; it amounts to the rejection of their philosophical positions. It is no more neutral and traditional than the doctrine of the ideality of time and space.

#### 4.2.2 Descartes and Locke

What about seventeenth century sources? There is no trace of the idea that logic is distinctively formal in Descartes or the British Empiricists. In fact, both Descartes and Locke articulate their conceptions of logic in explicit *opposition* to a Scholastic concern with schematic rules or “forms” of argument. For example, in the *Regulae*, Descartes represents his target as the view that logical rules

... prescribe certain *forms of reasoning* in which the conclusions follow with such irresistible necessity that if our reason relies on them, even though it takes, as it were, a rest from considering a particular inference clearly and attentively, it can nevertheless draw a conclusion which is certain simply *in virtue of the form*. (R:405-6, emphasis added)

Descartes and Locke offer two main criticisms of this conception of logic.<sup>28</sup> First, we are more likely to be taken by sophisms if we attend only to form than if we think directly about the ideas relevant in the particular case with our “untrammelled reason”:

<sup>27</sup>Beiser 1987 explains that “... the Wolffians maintain that the principles of logic are not ‘laws of thought’, but laws of all being; and these laws are true for ‘things-in-general’, whether these things be appearances or things-in-themselves, noumena or phenomena. ... Since both thought and reality have to conform to the laws of logic, we can rest assured that thought conforms to reality (and conversely); for both concept and object share a common logical structure” (201).

<sup>28</sup>See Gaukroger 1989 for the intellectual background of these criticisms.

But, as we have noticed, truth often slips through these fetters [i.e., the formal rules], while those who employ them are left entrapped in them. . . . Our principal concern here is thus to guard against our reason's taking a holiday while we are investigating the truth about some issue; so we reject the forms of reasoning just described as being inimical to our project. (Descartes R:406; cf. R:372-3, 439-40, Locke E:IV.xvii.§4, 670, 676, 678)<sup>29</sup>

Second, the syllogistic rules cannot lead us to new knowledge:

But to make it even clearer that the aforementioned art of reasoning contributes nothing whatever to knowledge of the truth, we should realize that, on the basis of their method, dialecticians are unable to formulate a syllogism with a true conclusion unless they are already in possession of the substance of the conclusion, i.e. unless they have previous knowledge of the very truth deduced in the syllogism. It is obvious therefore that they themselves can learn nothing new from such forms of reasoning. . . . Its sole advantage is that it sometimes enables us to explain to others arguments which are already known. (R:406; cf. Locke E:IV.xvii.§4, 6723, §6)

In place of the old Scholastic logic, Descartes and Locke articulate a new conception of logic as a set of rules for the correct use of our mental faculties, grounded in a kind of normative psychology of the mind.<sup>30</sup> Unlike syllogistic, this new logic is supposed to be a tool for the extension of our knowledge. Thus, in urging his students to start with the study of logic, Descartes adds:

I do not mean the logic of the Schools, for this is strictly speaking nothing but a dialectic which teaches ways of expounding to others what one already knows or even of holding forth without judgment about things one does not know. Such logic corrupts good sense rather than increasing it. I mean instead the *kind of logic* which teaches us to *direct our reason with a view to discovering the truths*

<sup>29</sup>Locke asks, "Of what use then are *Syllogisms*? I answer, Their chief and main use is in the Schools, where Men are allowed without Shame to deny the Agreement of *Ideas*, that do manifestly agree. . . ." (E:IV.xvii.§4, 675).

<sup>30</sup>Buickerood 1985 describes the new "facultative" conception of logic as follows: "The end of logic was taken to be the depiction of the natural history of the understanding on the basis of the assumption that the possibilities for the success of logic as a normative discipline hinged upon the possession of accurate and complete descriptions of those cognitive operations upon which prescriptions were to be levied" (187). As logic comes to be seen as a kind of normative psychology—the study of the proper functioning of the cognitive faculties—Locke's *Essay* begins to be regarded as a work of *logic* (Buickerood 1985:183; cf. section 4.3.3, below). That Hume regards Book I of the *Treatise* as logic is clear from his claim that "The sole end of logic is to explain the principles and operations of our reasoning faculty, and the nature of our ideas. . ." (xv, cf. 175). For the background of this conception in late Scholastic logic, see Gaukroger 1989:46-7 and Normore 1993.

*of which we are ignorant.* (Preface to French edition of the *Principles*, P:13-14, emphasis added; cf. Buickerood 1985:179)

This new logic does not offer schematic inference patterns, but heuristic rules for avoiding error and solving problems, most prominently: “We never go wrong when we assent only to what we clearly and distinctly perceive” (§43). These rules require *attention* to the particular contents with which the inference is concerned, not *abstraction* from them.

In neither Descartes nor Locke is there a distinction in kind between *formal* knowledge of logical relations and *material* knowledge of the truth of propositions. For both thinkers, knowledge of inferential connections consists in the intuitive perception of the relations between things or Ideas, just like knowledge of the premises themselves (Descartes R:440, Locke E:IV.xvii.§15). Hence, as Buickerood 1985 argues, “. . . on the view of neither Descartes nor Locke is any licit distinction between the form and the content of reasoning obtainable” (180):

. . . formal properties were not understood to account for the success of putative conceptual connections. It was the cognitive agent's perception of the ‘agreement or disagreement of ideas’, an event in the natural history of the understanding in which form and content are indissoluble, that ultimately grounded the derivations of truths. (188)<sup>31</sup>

It is not surprising, then, that one finds little talk of “form” in the *Port-Royal Logic*, “the most influential logic from Aristotle to the end of the nineteenth century” (Arnauld and Nicole 1996:xxiii). All one finds are the distinction between formal and material errors in argument (Arnauld and Nicole 1662:218) and an admonishment to “examine the soundness of an argument by the natural light rather than the forms of reasoning” (205). If there is an antecedent for Kant's doctrine of the formality of logic, it is not to be found in this tradition.

---

<sup>31</sup>Cf. Belaval 1960: “. . . pour Descartes, la raison n'a pas de structure” (64).

### 4.2.3 Leibniz

Leibniz is a more plausible candidate source for logical hylomorphism, for three reasons. First, he defends “the art of conducting arguments formally” against the Cartesian and Lockean attacks. Second, he urges the development of a science of combinatorics, more general than algebra, which he calls a “science of forms.” Third, he is the first logician to treat logic *algebraically*, setting down mechanical rules for the manipulation of logical formulas, without regard to their meaning. I will argue, however, that when Leibniz talks of “formality” in connection with logic, he means nothing more than *schematic formality*. Although Leibniz *does* think of logic as constitutive of thought as such (1-formal), he does not use the word “formal” to mark this feature. Most important, he does not think of logic as *3-formal*. Logic, for Leibniz, is a contentful science capable of extending knowledge: the norms governing thought as such are equally the most general laws governing *things*.

#### Arguing in form

Leibniz's *New Essays* are a good place to start, because we know that Kant read them in 1769.<sup>32</sup> There he would have found Leibniz defending scholastic logic against Locke's attacks, arguing that far from being useless, “. . . nothing could be more important than the art of conducting arguments formally, in accordance with true logic” (NE:IV.xvii.§4, 482-3). Leibniz clearly distinguishes formal correctness (i.e., the validity of inferences) from the correctness of the content (i.e., the truth and clarity of the premises) and implies that *logic* concerns only the first:

You can see that this reasoning is a sequence of syllogisms which is wholly in conformity with logic. I do not want now to go into its content, about which there are perhaps some things to be said or clarifications to be asked for. . . (482-3)

His only *criticism* of scholastic logic is that it is not sufficiently general to deal with all types of reasoning:

---

<sup>32</sup>See Remnant and Bennett's preface, NE:xiii.

But it must be grasped that by 'formal arguments' I mean not only the scholastic manner of arguing which they use in the colleges, but also any reasoning in which the conclusion is revealed *by virtue of the form*, with no need for anything to be added. (478, emphasis added; cf. L:389)<sup>33</sup>

If we think of Kant's logical hylomorphism as a traditional conception of logic, we will quite naturally read it into these passages from Leibniz. Like Kant, we will imagine, Leibniz thinks of logic as concerned with the mere *form* of thought, in abstraction from its relation to objects and the world. From this perspective, the most salient difference is that Leibniz sees the limitations of traditional syllogistic to which Kant, the inferior logician, is blind.

But there is no warrant for reading Kant back into Leibniz here. When Leibniz says that an inference concludes "by virtue of its form," he means only that it is an instance of a general, schematic rule of inference, not that it abstracts from all relation to objects and the world. Thus, where Kant invokes formality to distinguish pure general logic from mathematics and from the "special logics" that provide the principles of particular sciences, Leibniz argues that the scholastic inference forms must be *extended* to include specifically mathematical and scientific forms of argument. Even algebra, infinitesimal analysis, and the Euclidean rules for converting ratios count as techniques of formal argument, in Leibniz's sense, because "in each of them the form of reasoning has been demonstrated in advance so that one is sure of not going wrong with it" (NE:479).<sup>34</sup>

To be sure, Leibniz makes a distinction corresponding to Kant's distinction between general and special logics. Whereas general logic is not tied to any particular subject matter, the forms of the special logics are restricted to a particular domain (e.g., quantity):

Euclid's invertings, compoundings and dividings of ratios are merely particular kinds of argument form which are *special* to the mathematicians and to their subject matter; and they demonstrate [the soundness of] these forms with the aid of the universal forms of *general* logic. (NE:479)<sup>35</sup>

<sup>33</sup>Cf. L:389, 465-6, 479; GP V.7, quoted in Couturat 1901:101 n. 4.

<sup>34</sup>At NE:523, Leibniz even talks of "the logic of medicine, which is concerned with the art of finding methods of cure..."

<sup>35</sup>Leibniz holds that all valid forms of argument can be demonstrated from a few basic ones: non-syllogistic inferences "... are nevertheless demonstrable through truths on which ordinary syllogisms themselves depend" (NE:479). But in his commentary on Descartes' *Principles of Philosophy*,

But for Leibniz, the distinction between general and special logics does not coincide with the distinction between the formal and the non-formal, as it does for Kant. Both general and special logics are formal in the only sense countenanced by Leibniz: that is, *schematically* formal. Thus there is no reason to read any interesting doctrine of logical hylomorphism into Leibniz's talk of formal argument.

### The “science of forms”

Aware of the logical limitations of traditional syllogistic, Leibniz devotes much effort to the development of a “universal characteristic” that resembles algebra in that its rules hold no matter what the signs stand for (e.g. GP:VII.221)—abstracting, we might be tempted to say, from the *content* of thought. Since Leibniz conceives of all reasoning as connection and substitution of characters (GP:VII.31; VII.204; Couturat 1901:102), the operations of his universal characteristic might naturally be taken to represent forms *of thought*. In fact, Leibniz characterizes his “universal characteristic” as a “combinatoric art,” which he describes as a “science of forms” (GM:IV.451, quoted Couturat 1901:287 n. 2).<sup>36</sup> This combinatoric art plays the role given in the *New Essays* to the “general forms of logic”: the special forms of algebra and the other mathematical sciences are all regarded as applications of the *ars combinatoria* to particular domains (e.g., magnitudes and indeterminate numbers) (Couturat 1901:289-90; cf. GP IV.295-6, L:250).<sup>37</sup>

Again, if one is not alive to the possibility that the Kantian conception of logic is highly original, it is easy to read it into Leibniz. In his pioneering study of Leibniz's logic, Couturat gives a Kantian interpretation to nearly all of Leibniz's references to form, often quite Leibniz makes it clear that this demonstration requires that “the peculiar nature of the subject [be] taken into consideration” (L:391).

<sup>36</sup>“But for me the *art of combinations* is in fact something far different, namely, the science of forms or of similarity and dissimilarity, while algebra is the science of magnitude or of equality and inequality. The combinatory art seems little different, indeed, from the general science of characteristics, by the use of which fitting characters have been or can be devised for algebra, for music, and even for logic itself” (trans. L:192). For the identification of general combinatorics and the universal characteristic, see also Couturat 1901:286.

<sup>37</sup>For a useful discussion of the relation of logic and the universal characteristic, see Couturat 1901:299 n. 3,4.

inappropriately. For example, he says that in characterizing combinatorics as a “science of forms,” Leibniz intends “not only mathematical formulas and algebraic ‘forms’, but all the forms of thought, that is to say the general laws of the mind” (1901:299). But this reference to “forms of thought” is unwarranted: in the passages Couturat quotes, Leibniz is using “form” to mean “quality”, as opposed to the object of algebra: quantity.<sup>38</sup> The *ars combinatoria* is more general than algebra because it is not restricted to *quantitative* features of objects, which can be equal or unequal, but concerns also *qualitative* features, which can be similar or dissimilar. Thus the “science of forms” is meant to apply to features like *shape*, not to “forms of thought.” In fact, although Leibniz writes elsewhere of “forms of argument,” he does not use the phrase “forms of thought” in any of the passages Couturat quotes. Indeed, Leibniz says that combinatorics concerns the forms of *things*.<sup>39</sup>

### The algebraic treatment of logic

Leibniz was the first philosopher to see the possibility of constructing an abstract logical calculus, governed by precisely specified syntactic rules and admitting of any interpretation consistent with these rules. After laying down rules for a sign “ $\oplus$ ,” including “ $B \oplus N = N \oplus B$ ” and “ $A \oplus A = A$ ,” he notes that

... wherever these laws just mentioned can be used, the present calculus can be applied. It is obvious that it can be used in the *composition* of absolute concepts, where neither laws of order nor laws of repetition obtain... The same thing is true when certain given things are said to be *contained* in certain things... (GP:VII.236 ff., trans. Kneale and Kneale 1964:343, emphasis added)

That is, the same axioms hold whether  $\oplus$  is interpreted as *conjunction* (in which case  $A \oplus B = C$  indicates that the concept  $C$  is formed by the composition of  $A$  and  $B$ ) or as *disjunction* (in which case  $A \oplus B = C$  indicates that anything in the class  $C$  is either in the

<sup>38</sup>Cf. GP:VII.B vi,9, quoted in Couturat 291 n. 2: “Imaginatio generaliter circa duo versatur, Qualitatem et Quantitatem, sive magnitudinem et formam; secundum quae res dicuntur similes aut dissimiles, aequales aut inaequales” (“Imagination in general considers two things, quality and quantity, or magnitude and form; according to which things are said to be similar or dissimilar, equal or unequal”).

<sup>39</sup>“... id est de formis rerum, abstrahendo tamen animum a magnitudine, situ, actione” (GP:V.7, 4, quoted by Couturat at 288 n. 2).

class A *or* in the class B). Leibniz notes further that the same laws that govern the inclusion relations of concepts in logic can be applied in other areas, e.g. in the study of coincident lines in geometry (Couturat 1901:305). These observations mark an important conceptual advance, one we associate with the *formalization* of logic and mathematics. Kneale and Kneale remark:

We have discussed the possibility of different interpretations because the most interesting feature of Leibniz's calculus is its abstract *formality*. . . . Here for the first time there was an attempt to work out a piece of abstract mathematics, i.e. mathematics that is not specifically concerned with space or numbers. (344-5, emphasis added)

Similarly, Couturat writes:

One sees that Leibniz has outlined a general theory of operations, considered in their properties and their formal relations, and that he has already had the entirely modern idea of considering the algebraic signs themselves as symbols of indeterminate operations. One understands also why this logic of relations is at the same time a combinatorics and a characteristic: it is because the relations of objects are expressed by the *entirely formal combinations of signs* that represent these objects, and because these relations themselves, studied in their *form*, are represented by symbols of variable or indeterminate sense. (1901:303, first emphasis added; cf. 319, 101)<sup>40</sup>

However, the sense in which Leibniz's logical calculus is "formal" is not one that supports logical hylomorphism. Logical hylomorphism is the view that logic can be *distinguished* from the other disciplines by its formality. To give a formal *treatment* of logic is one thing; to say that logic itself is distinctively formal is quite another. Leibniz's logical calculus is admittedly *syntactic-formal*. But as we saw in section 2.1, syntactic formality is inadequate by itself for delineating logic. It does not preclude relation to objects or to a subject matter. Granted, if the symbols " $\oplus$ " and " $=$ " are considered as purely syntactic objects, then they have no relation to objects, but so considered, they have nothing to do with logic, either. To regard a syntactically defined system as a *logic* is to interpret at least some of its terms, and the fact that syntactic rules can be given for interpreted terms shows nothing about these terms' relation to objects or the world.

---

<sup>40</sup>All translations from Couturat are my own.

At any rate, Kant would not have thought of Leibniz's logical calculi as "formal" in his sense, since they involve the manipulation of sensible *symbols*. An important advantage of the universal characteristic, Leibniz holds, is that it allows us to bring imagination and sense to bear on abstract reasoning: as Couturat says, "...the characteristic translates thought into an intuitive form.... The abstract laws of logic are thus translated by the intuitive rules which govern the manipulation of signs" (89). The point of the universal characteristic is to give to reasoning in metaphysics the help which imagination and intuition provide in mathematics (93). Indeed, Kant makes a similar point about algebra in the first *Critique*:

Even the method of algebra with its equations, from which the correct answer, together with its proof, is deduced by reduction, is not indeed geometrical in nature, but is still a characteristic construction.<sup>41</sup> The concepts attached to the symbols, especially concerning the relations of magnitudes, are presented in intuition; and this method, in addition to its heuristic advantages, secures all inferences against error by setting each one before our eyes. (A734/B762; cf. A717-B745 on "symbolic" vs. "ostensive" construction)

But for Kant, algebra's essential use of intuition is enough to make it non-formal. So Leibniz's algebraic treatment of logic would not have led Kant to his logical hylomorphism, even had he known about it.<sup>42</sup>

### Logic as an instrument of discovery

Indeed, although Leibniz's logic is schematic and makes use of purely syntactic rules, it does *not* abstract from all content or relation to objects. To the contrary: Leibniz thinks of logic

---

<sup>41</sup>*charakteristische Konstruktion*: here I depart from Kemp-Smith, who has "is still constructive in a way characteristic of the science."

<sup>42</sup>Manley Thompson 1972-3 raises a related problem about Kant's own logic. If general logic "contains symbolic construction and demonstrations," he says, "it would seem at least in this sense to be something Kant would have to regard as a branch of mathematics. One may be tempted to take Kant's reconstructed position to be that general (formal) logic uses symbolic constructions and demonstrations to determine valid forms of discursive proof, and is thus a special case of math. This is essentially the position of C. S. Peirce. But this position conflicts with Kant's view that logical possibility is purely conceptual and constructibility intuitive" (342 n. 23). What this shows, I think, is that Kant does *not* think of general logic as containing symbolic construction and demonstration. But as Thompson notes, Kant does use constructions in his general logic: "...the square of opposition, the use of circles and squares to diagram the relation of S and P in categorical judgments (*cf.* *Logic* §§21, 29), and the four figures of the syllogism...."

as a tool for the discovery of new truths: “By logic or the art of reasoning I understand the art of using the understanding not only to judge proposed truth but also to discover hidden truth” (GP VII.514-27 = L:463). The result is a blurring of the line between metaphysics and logic:

...the true metaphysics is hardly different from the true logic, that is to say, from the art of invention in general. (Letter to Duchess Sophie, GP IV.291-2, my translation)

The true logic is not only an *organon*, but also in a certain way comprises principles and the true method of philosophizing, because it treats of those general rules, by which truths and falsehoods can be distinguished. (GP IV.137, my translation)<sup>43</sup>

Faced with these passages and others like them, even Couturat is forced to admit that “... Leibniz considers logic as a *real* science, not simply a *formal* one” (279 n. 4). For as we have seen, a truly *formal* logic (in Kant's sense) would be incapable of serving as an organon capable of extending knowledge: because it abstracts from all relation to objects, it cannot tell us anything about them (A60-1/B85). Kant's remark in the *Jäsche Logic* that “[l]ogic is thus not a universal art of discovery, to be sure, and not an organon of truth—not an algebra, with whose help hidden truths can be discovered” (JL:20) is surely aimed directly at Leibniz's ideal of a universal characteristic.

It follows from Leibniz's conception of logic as an instrument of discovery that it must be concerned with relations of *real* compatibility between concepts. Descartes' ontological proof of the existence of God is flawed, according to Leibniz, because it assumes that “a being with all perfections” is a possible concept (GP IV.294, GP VII.292-8 = L:229-233). Before using logic to infer that a being with all perfections has the perfection *existence*, one must make certain that the logical “matter” involved in this inference—the concept of a being with all perfections—is not contradictory. Thus there is a *presupposition* for the application of the universal characteristic, for the same logical moves that verify truths will verify falsehoods (like “the Pontiac with all perfections exists”) when applied to impossible

<sup>43</sup>“Logicam veram non tantum instrumentum esse, sed et quodam modo principia ac veram philosophandi rationem continere, quia generales illas regulas tradit, ex quibus vera falsaque judicari... possunt.”

concepts. As Couturat put it: "...one can never deduce with security from an arbitrary definition, unless one knows that the object defined is possible, i.e., non-contradictory: because if it were impossible (contradictory), one could deduce mutually contradictory consequences from its definition" (189).<sup>44</sup>

If logic is to extend knowledge, then, it must concern itself with the "matter" of thought as well as its form. This is why the universal characteristic presupposes the elaboration of an "encyclopedia" which gives the analyses of the basic concepts of all knowledge into their simplest constituents (Couturat 117). Leibniz says of the characteristic that "whoever learns this language at the same time learns an encyclopedia" (Letter to Oldenburg, GP:VII.13, quoted in Couturat 100 n. 1, my translation). It is for this reason that Trendelenburg 1870 denies that Leibniz's *ars combinatoria* is a formal logic ("of which he knew nothing"): Leibniz brings in "the real disciplines" to sort the valid from the invalid combinations of concepts (23).<sup>45</sup>

For Kant, by contrast, logic looks *only* at how concepts have been linked together by the understanding; it abstracts entirely from the particular contents of these concepts. Thus, it does not need to worry about the real possibility of the things designated by the linked-together concepts: it can draw all the consequences it wants from "x is a round square." "Anything we please can be made to serve as a logical predicate; the subject can even be predicated of itself; for logic abstracts from all content" (A598/B626). It is not *logic's* task to tell us that "the round square is round" is not true (because it relates to no object): for that, we need geometry, which takes us outside of the mere form of thought and into the realm of pure intuition.<sup>46</sup> Similarly, the problem with the ontological argument is *not* that

<sup>44</sup>In "A Specimen of Universal Calculus" (GP VII.218-27; L:241), Leibniz writes: "It makes no difference if the terms sometimes combined in this way are inconsistent. Thus a circle is a null-angle. A square is a quadrangle. Therefore a square circle is a null-angled quadrangle. For the proposition is valid, though from an impossible hypothesis." Here he seems to deny that the application of logic has any presupposition. But he will have to admit that logic acquires a presupposition whenever it is put to use as an instrument of discovery.

<sup>45</sup>On similar grounds, Russell 1900 argues that logic for Leibniz is *synthetic* (since it presupposes synthetic judgments about the compatibility of concepts, §11).

<sup>46</sup>Although Kant holds that the denial of an analytic truth can be shown to be logically contradictory (A151/B190), he is *not* committed to the converse view that any (purported) judgment whose denial is logically contradictory is an analytic truth. For the subject concept may fail to have

the concept of a supreme being might be inconsistent or impossible—that would not stop logic from applying to it and deriving the existence claim—but that through concepts alone we can never show the existence of an object (A601/B629).

### Leibniz's formalism

It is one thing to recognize the usefulness of “forms of argument” and recommend the development of an algebraic “universal characteristic,” quite another to see logic as a discipline that concerns only the form of thought, in a sense that excludes concern with content. Though Leibniz makes the first move, he does not make the second, and his view that logic can be an instrument of discovery is plainly incompatible with it. As Belaval 1960 notes, “. . . formalism, for Leibniz, is not empty, it is full of being: in stating the principles of universal Reason, it states, at the same time, the principles of the World in which this Reason is inscribed” (34-5).

#### 4.2.4 Scholastics

I will not concern myself here with the question of whether Kant was influenced by the *medieval* distinction between formal and material consequence. Despite the virtual disappearance of medieval logic in the wake of the Humanists' critiques (see Ashworth 1982, Normore 1993), Kant may have been familiar with some versions of the medieval distinction. But he never discusses the medievals in connection with his delineation of logic, nor does he accord them much importance in any of the capsule histories of logic in the lectures (see section 4.3.3, below). This would be inexplicable if he took the idea that logic (or an important branch thereof) is distinctively formal from scholastic sources.

Indeed, it is a long way from the characterization of formal consequence one finds in, say, Buridan, to the Kantian notion of formality as abstraction from all content.<sup>47</sup> Formal an object, and truth is, for Kant, agreement of knowledge with its object (A58/B82). Logic alone can never determine whether a concept determines an object. This is why Kant allows analytic propositions in mathematics “only because they can be exhibited in intuition” (B17). See Beck 1955 for a useful contrast of Kant's views on analyticity with those of the Leibnizians.

<sup>47</sup>For Buridan's account of formal consequence, see page 256, below.

consequences, as Buridan understands them, hold for all uniform substitutions of *categoric* terms, but they are not indifferent to changes in the *syncategoric* terms. In appendix A, I suggest that an argument from the *schematic* formality of the syllogism to its *3-formality* can be found in Abelard. But Abelard's argument presupposes characteristically medieval views about ontology. Abelard can reject the view that the validity of syllogisms depends on "the nature of things" only because he does not think that facts of the form

A, B, and C stand in the relation Q, or  
 not both: { (all A are B and all B are C) and not (all A are C) }, or even  
 for all A, B, and C: A, B, and C stand in the relation Q,

can be facts about "the nature of things." Kant does not share this supposition: he takes seriously the fact that the laws of Newtonian physics cannot be forced into the mold of substance-attribute predications but essentially involve relations between multiple entities.<sup>48</sup> The "*nature of things* (that is, of things as appearances)" (KrV:A228/B281) is a relational nexus. So Kant cannot use Abelard's argument to show that syllogistic validity does not depend on facts about objects in the world.<sup>49</sup> Even if (as seems unlikely) Kant got his hylomorphic *terminology* from the medievals, his *motivation* and *articulation* of the idea that logic is distinctively formal would have to be counted as wholly novel.

A more promising medieval antecedent for Kant's logical hylomorphism would be the widely-held scholastic view that logic is distinguished from metaphysics by its concern with *second intentions*: that is, with "beings of reason"—e.g., genus, species, subject, predicate, syllogism—which are not found in "the nature of things" but contrived by reason

---

<sup>48</sup>See the Introduction to Friedman 1992 for an account of how Kant's departures from Leibnizianism were motivated by a desire to take Newtonian physics seriously. "For Leibniz and Wolff space is ideal because relations between substances are ideal. . . . For Kant, by contrast, relations of interaction between substances are in no way ideal: a universal principle of mutual interaction is a distinct reality over and above the mere existence of substances. . ." (7-8).

<sup>49</sup>I suggest in appendix A that Abelard's argument has the (perhaps unintended) consequence that a "relational syllogism" like "A is to the east of B; B is to the east of C; therefore, A is to the east of C" is, like categorical syllogisms, good in virtue of its structure or form. But Kant would certainly not regard this inference as formally valid, since its validity depends on the structure of space.

in considering things (Bochenski 1956:26.04-26.08, Schmidt 1966:53, 306-8). Sir William Hamilton found this view to be a mere “variant” on the view that logic is concerned with the form of thought, to the exclusion of its matter (1867:20; cf. Thomson 1860:24-5), and some modern Thomists have suggested that second intentions can be thought of as logical forms (Simmons 1961:63-4, Schmidt 1966:69-70). Flynn 1946 even suggests that “...Kant seems to have been quite confident that his own conception of Logic was the traditional one” (though he argues that in fact “it implies a notion of the nature of Logic which is contrary to Aristotelian and Thomistic teaching”).

However, there is no textual evidence that Kant's delineation of logic was influenced by the scholastic view that logic concerns second intentions. If Kant were self-consciously returning to an earlier tradition, against the dominant current of modern philosophy, one would expect him to acknowledge this. Moreover, there are significant differences between the medieval second intentions and Kant's forms of thought. As Schmidt 1966 argues,

... we must not conclude from the fact that logical intentions exist only in the mind that they are pure forms of the intellect without any content or that logic is “formal” in the sense that it takes no account of the things that are known. If the distinction were forced upon us we should have to say that for St. Thomas logic is rather “material” than “formal,” since it necessarily takes into account the natures of the things known. It is only because the nature of *some thing* is in the intellect that logical intentions of it are formed; and these, as accidents of the nature, can no more be considered without reference to their subject than a real accident can. (308)

While Kant's logical forms are in some sense *prior* to real objects, Aquinas's second intentions are *posterior*.<sup>50</sup>

---

<sup>50</sup>Cf. Flynn 1946: “Second intentions are relations which are formed by the mind through comparison of objects and which, therefore, have their foundations in first intentions—in known objects: ‘relationes quae attribuuntur ab intellectu rebus intellectis, prout sunt intellectae’, as St. Thomas explains. Since relations are known only through their foundations, it is impossible for any part of Logic to treat of forms which have no reference to what is now usually called the content of thought” (181).

### 4.2.5 Crusius, Lambert, Tetens

It would be misleading to portray Kant's delineation of logic as formal as *entirely* novel: we can find anticipations in Crusius, Lambert, and Tetens, all of whom influenced Kant. Crusius distinguishes between logical and real reason and urges the need for material principles in metaphysics as well as formal ones (de Vleeschauwer 1939:12, Friedman 1992:22). Similarly, Lambert, in a 1765 letter to Kant, distinguishes between "principles" that derive from "the *form* of knowledge" and "axioms" that derive from "the *matter* of knowledge":

We do not get to any material knowledge from the form alone; we shall remain in the realm of the *ideal*, stuck in mere nomenclature, if we do not look out for what is primary and thinkable in itself, the matter or *objective* stuff of knowledge. (Nov. 13, 1765: PC:44-5, Ak:X.51-4)

This is clearly an anticipation of Kant's distinction between logic and the contentful sciences.<sup>51</sup> Finally, Tetens, whom Kant studied in the last few years before publishing the first *Critique*, "...distinguishes between matter and form in knowledge" and (like Kant) associates the formal with the subjective (de Vleeschauwer 1939:69).

These glimmerings of logical hylomorphism deserve further study.<sup>52</sup> But they do not contradict the fundamental thesis I am urging: that modern logical hylomorphism and transcendental idealism are, as it were, fraternal twins. My claim, recall, was not that Kant's logical hylomorphism was entirely novel, but that it has just as much claim to being called a Kantian invention as transcendental idealism. That we find anticipations of logical hylomorphism in three philosophers that are generally regarded to have been important influences on the development of Kant's transcendental idealism (de Vleeschauwer 1939, Friedman 1992) only supports my thesis. Twins are born together, but they also share the same womb and the same gestation.

<sup>51</sup>Cf. Longuenesse 1998:150 n. 26: "This internalization within thought of the relation between *matter* and *form* is not, on the whole, a complete novelty. . . . In his *Architectonic*, Lambert deals with the 'logical form,' 'derived from the *operations of the understanding* alone.'" See Lambert A:XXII, A:39, de Vleeschauwer 1939:45.

<sup>52</sup>And they are only glimmerings. For example, Lambert puts mathematics together with logic on the "formal" side; as for Crusius, see JL:21, quoted in section 4.3.3, below.

### 4.3 The genesis of logical hylomorphism in Kant

The evidence reviewed so far makes it plausible that the delineation of logic by its formality is Kant's own invention. Far from being something he takes over from his modern predecessors, logical hylomorphism goes *against* the tradition (in both its rationalist and empiricist branches). I now want to suggest that if we compare Kant's lectures on logic from different periods, we can actually *see* this conception of logic being invented. Kant does not begin describing logic as concerned exclusively with form until the critical period, around the time he is writing the first *Critique*. In this section, I will examine the difference in treatment between the early, pre-critical logic lectures (*Blomberg Logic*, *Phillipi Logic*)<sup>53</sup> and the later, critical logic lectures (*Vienna Logic*, *Pölitz Logic*, *Busolt Logic*, *Dohna-Wundlacken Logic*, *Jäsche Logic*)<sup>54</sup> with respect to three related topics: (1) the definition of logic, (2) whether logic offers a material criterion of truth, (3) the history of logic.

#### 4.3.1 The definition of logic

The later logic lectures all characterize logic as concerned with the form of thought, abstracting from content (DWL:693-4, VL:791, JL:12, BuL:609, PzL:503). On the other hand, the earlier lectures follow Meier's definition of logic closely and say nothing about formality. Admittedly, the *Blomberg Logic* does distinguish between the formal and the material in cognition (i.e., between the manner of representation and the object, BL:40; cf. PhL:341). Here Kant goes beyond the passage of Meier on which he is commenting (§11-12), which merely distinguishes the cognition from its object. But Kant goes on to say: "Logic has to do *for the most part* with the formal in cognition" (my emphasis), employing a qualification that can have no place in his mature conception of logic. The *Phillipi Logic*, which distinguishes the matter and form in cognition in much the same way as the *Blomberg Logic*, also

<sup>53</sup>These date from the early 1770s—after the *Dissertation*, but before Kant has given up on purely conceptual knowledge of things in themselves (see de Vleeshauwer 1939:62-6). Translations from the *Blomberg Logic* are from Kant 1992, those from the *Phillipi Logic* are my own.

<sup>54</sup>These date from around 1780 to 1800. Translations from the *Vienna*, *Dohna-Wundlacken*, and *Jäsche Logics* are from Kant 1992; those from the *Pölitz* and *Busolt Logics* are my own.

fails to regard formality as distinctive of *logic*: “All philosophy treats only of form” (341)!<sup>55</sup>

The closest we come to Kant’s mature conception of logic in the early lectures is in the *Phillipi Logic*: “In logic we consider not the determinations and relations of things, but rather only the determinations and relations of concepts and their relations” (339). Since the thought here appears very close to the thought that logic concerns only the *form* of thought, abstracting from its objects, it may seem surprising that Kant does not here use the hylomorphic terminology. But there is a good reason for this. For Kant goes on to say that arithmetic, too, “. . . has no real object, but only teaches about the concepts of numbers and their relations” (339-340). Arithmetic is *like* logic in being concerned with concepts rather than things; the difference between them is only that logic is not limited to the particular concepts of “numbers and their relations,” and thus has even greater generality. There is no hint of the chasm that we find in Kant’s critical work between the contentful science of arithmetic and the purely formal science of logic (see section 4.1.2, above).

Nor, finally, is formality invoked in distinguishing metaphysics from logic. Both sciences, Kant says, are concerned with objective laws of reason (PhL:313). The difference is that metaphysics is concerned “only with *pure* reason, which is not mixed with sensibility and takes its principles entirely from reason and not from experience” (313), while logic “borrows its principles in part from reason, in part from experience” (314). For the critical Kant, by contrast, it is logic that concerns the laws of reason and the understanding by themselves; metaphysics must consider them in conjunction with sensibility, for otherwise it can say nothing about objects.

### 4.3.2 Material criterion of truth

In the first *Critique*, the *Jäsche Logic*, and the other lectures from the critical period, Kant denies that a universal material criterion of truth is possible, on the grounds that it would simultaneously have to be sensitive to and abstract completely from differences

---

<sup>55</sup>Compare GMS:387: “Formal philosophy is called logic. Material philosophy, however, has to do with determinate objects and with the laws to which these objects are subject. . . .”

among objects (JL:50, VL:823, DWL:718-19, VL:823, KrV:A58-60/B83-4). Hence logic can provide only a formal criterion of truth—a *necessary* (but insufficient) condition for the truth of a cognition, the mere agreement of the cognition with itself (JL:51). In the framework of transcendental idealism, this agreement of the understanding with itself is not enough to secure real truth: relation to an object must also be secured, hence the faculty of sensibility must be in play.

Hence, what Kant calls “logical truth” in the later works (e.g. at JL:51 and R:2145) is not what we mean by that phrase, but rather consistency and groundedness (the existence of grounds and the absence of false consequences). Logical truth is *formal* truth, which “. . . consists merely in the agreement of cognition with itself, in complete abstraction from all objects whatsoever and from all differences among them” (JL:51). In the *Blomberg Logic*, on the other hand, “logical truth” means *factual* truth. When Kant says that logical truths “relate merely to the understanding and reason,” he is contrasting them not with truths that relate also to intuitive or sensible content, but rather with truths that relate to the “condition of taste” (aesthetic truth) or “the rules of free will” (practical truth) (91-2). Thus, when he says that fables seldom contain logical truth (92), he does not mean that they are seldom logically consistent. He means that they describe states of affairs that are not factually true, like animals talking. In both early and later lectures, then, “logical truth” is characterized as “relating merely to the understanding and reason,” but it is only in the later lectures that this characterization is taken to imply that logical truth is merely formal.

To say that logic can offer no general material criterion of truth is to say that it cannot serve as an organon for the extension of knowledge, but only as a canon. The critical logics all deny that logic is an organon on precisely these grounds (JL:13, DWL:695, VL:792, BuL:610, PzL:505).<sup>56</sup> The *Phillipi Logic*, on the other hand, explicitly states that logic is

<sup>56</sup>Some care is necessary here, because Kant still allows that general logic, “considered as practical,” can be an “organon of scholastic method” (JL:18). What he means is that logic can lead to a *formal* extension or articulation of knowledge (cf. JL:64). What general logic *cannot* be is an organon of the sciences.

an organon of the sciences (314, 317). In the *Blomberg Logic*, too, we find logic put on the same footing as the real sciences, in a way that would be anathema to the later Kant: “Our rules have to be governed by those universal basic truths of human cognition that are dealt with by *ontologia*. These basic truths are the *principia* of all sciences, consequently of logic too” (28).<sup>57</sup>

If to say that logic is formal implies that it is independent of material truth in a way that disqualifies it from being an organon, then Kant’s early lectures preclude this conception of logic.

### 4.3.3 Remarks on the history of logic

Kant’s changing perspective on the history of logic are also revealing. In the *Jäsche Logic*, we read:

Among modern philosophers there are two who have set universal [i.e., general] logic in motion: *Leibniz* and *Wolff*. *Malebranche* and *Locke* did not treat of real logic, since they also deal with the content of cognition and with the origin of concepts . . . .

*Crusius* also belongs to the modern logicians, but he did not consider how things stand with logic. For his logic contains metaphysical principles and so to this extent oversteps the limits of this science; besides, it puts forth a criterion of truth that cannot be a criterion, and hence to this extent gives free reign to all sorts of fantastic notions. (21, cf. PzL:509, DWL:701)

Here the criterion of formality is used to deny that certain “logicians” are really doing logic. In the *Blomberg Logic*, however, there is no comparable criticism of Locke as a logician. Instead, we find: “Locke’s book *de intellectu humano* is the ground of all true *logica*” (37). And Crusius is criticized not for his criterion of truth, but for being too difficult to understand (37, 39). Similarly, in the *Phillipi Logic*, Locke’s logic is characterized as “not dogmatic, but critical” (338). The formality criterion is never used in the earlier lectures to criticize logicians for straying from logic’s true charter.

<sup>57</sup>Cf. the passages from Wolff quoted in section 4.2.1, above.

#### 4.3.4 Evidence of the Reflexionen

It appears, then, that Kant's views on the nature of logic changed radically during the mid 1770s—the period in which he was writing the first *Critique*. To the extent that we can trust Adickes' dating of the marginalia from Kant's text of Meier, they corroborate this view.<sup>58</sup> Of the many notes on the first few sections of Meier, the earliest characterize logic in much the same way as does the *Blomberg Logic* (for instance, logic is often said to be an organon: R:1566, 1573), while the later ones closely resemble the discussion of general logic in the *Jäsche Logic*. The first references to form and content in characterizing logic occur in a long *Reflexion* which Adickes dates from the early 1760s to the mid-1770s:

(The use of the understanding according to the form or the content)...

(Logic as canon (analytic) or organon (dialectic); the latter cannot be dealt with universally, because it is a doctrine of the understanding not according to the form, but rather according to the content.) (R:1579)

This *Reflexion* contains at least two temporal strata of comments, and the sentences referring to the form/content distinction are all marked by Adickes as later interpolations. If we put them towards the end of Adickes' date range, they date from the period between the earlier logic lectures and the first *Critique*—which is just what we'd expect. Other early references to logical hylo-morphism date from the same period: in R:1603 (dated 1773-5), we find the claim that logic does not make distinctions among objects, in R:3039 (dated 1773-5) we find a distinction between the form and matter of judgments, and in R:2155 and 2162 (dated 1776-8) we find the claim that logic abstracts from the matter (object) of knowledge and concerns only the form (R:2155, 2162).

<sup>58</sup>These marginalia are collected in Kant Ak:XVI (all translations from the *Reflexionen* are my own). Adickes' dates are regarded by many as unreliable. For his methodology, see the introduction to volume XIV of the *Akademie* edition, esp. xxv-liv. Relevant *Reflexionen* from this volume include 1579, 1603, 1608, 1612, 1620, 1624, 1627, 1629, 1721, 1904, 2142, 2152, 2155, 2162, 2174, 2178, 2225, 2235, 2324, 2834, 2851, 2859, 2865, 2871, 2908, 2909, 2973, 3035, 3039, 3040, 3045, 3046, 3047, 3053, 3063, 3070, 3126, 3127, 3169, 3210, 3286. Also relevant are *Reflexionen* 3946 and 3949 from Kant Ak:XVII (marginalia on Baumgarten's *Metaphysica*).

There are only four passages expressing logical hylomorphism that Adickes dates before 1773-5 (R:1579, R:1721, 3035, 2865).<sup>59</sup> We have already discussed R:1579. In each of the other three cases, Adickes expresses uncertainty about the dating and gives 1773-5 as an alternative. Thus, all of the evidence from the *Reflexionen* is at least consistent with the hypothesis that Kant's logical hylomorphism dates from the period 1773-1775.

#### 4.3.5 Precritical works

This is not to deny that there are anticipations of logical hylomorphism in Kant's earlier, precritical works.<sup>60</sup> In the *Inaugural Dissertation* of 1770, Kant distinguishes between *real* and *logical* uses of the understanding:

By the first of these uses, the concepts themselves, whether of things or relations, *are given*, and this is the REAL USE. By the second use, the concepts, no matter whence they are given, are merely subordinated to each other. . . and compared with one another in accordance with the principle of contradiction, and this use is called the LOGICAL USE. (ID:§5, 393)

In the *Inquiry* of 1764, he applauds Crusius's distinction between formal and material principles of reason, emphasizing against the Wolffians that formal principles are not sufficient for knowledge.<sup>61</sup>

And *Crusius* is also right to criticise other schools of philosophy for ignoring these material principles and adhering merely to formal principles. For on their basis alone it really is not possible to prove anything at all. (D:295)

---

<sup>59</sup>In R:2834, Kant distinguishes between the matter and form of concepts, but says nothing about *logic*.

<sup>60</sup>These precritical roots have been emphasized by Allison (1973:54) and Longuenesse (1998:150 n. 26).

<sup>61</sup>The formal principles are the law of identity and the law of contradiction (294). The material principles are indemonstrable propositions, like "a body is compound," which predicate of a concept one of its immediate characteristic marks (295). Kant says that "[t]he *form* of every *affirmation* consists in something being represented as a characteristic mark of a thing, that is to say, as identical with the characteristic mark of a thing," and "[t]he *form* of every *negation* consists in something being represented as in conflict with a thing. . ." (294). Cf. the *Metaphysik Herder* (lecture notes dating from 1764): "*Form* is the way in which I ought to compare the subject and predicate. *Matter* is which predicates ought to be compared with the subject" (Ak:XXVIII, 1:8, cited by Longuenesse 1998:150 n. 26 as the first appearance of "the idea of logical form" in Kant).

And as early as the *Beweisgrund* of 1763, Kant distinguishes between the material element (*das Materiale*) and the formal element (*das Formale*) in inconceivability or impossibility (B:77), identifying the latter with the logical element (*das Logische*):

... in every possibility we must first distinguish the something which is thought, and then we must distinguish the agreement of what is thought in it with the law of contradiction. A triangle which has a right angle is in itself possible. The triangle and the right angle are the data or the material element in this possible thing. The agreement, however, of the one with the other, in accordance with the law of contradiction, is the formal element in possibility. I shall also call this latter the logical element in possibility, for the comparison of the predicates with their subjects, according to the rule of truth, is nothing other than a logical relation. (B:77-8)

These passages show movement away from an ontologized Wolffian conception of logic (see section 4.2.1, above) and towards Kant's mature conception of logic as concerned with the form of thought in abstraction from all content. But they are stages along the way, not the finished product. A rationalist might well distinguish between "formal" and "material" principles, taking them to be principles of both thought *and* being. But it is essential to Kant's mature logical hylomorphism that the form of thought not be confused with the form of being: as Longuenesse puts it,

The various ways in which we combine our concepts in judgments and syllogisms are not the more or less adequate expression of ways in which essential and accidental marks are combined in things, but merely the implementation of the rules proper to our discursive activity. (1998:10)

In adopting Crusius's distinction between formal and material principles and limiting logic to the former, Kant has taken a first step towards a deontologized logic. But he still has not taken the decisive second step: declaring that logic abstracts entirely from relation to the content of thought. That point does not become clear until Kant abandons the idea that knowledge is possible through concepts alone, without relation to sensibility.

Thus, the existence of these *germs* of the hylomorphic delineation of logic in precritical works should not lessen the interest of the co-emergence of the hylomorphic delineation with Kant's transcendental idealism. In section 4.4, we will consider some philosophical reasons

for the connection between these two doctrines.

#### 4.4 Transcendental idealism and the formality of logic

In the preceding sections, I have argued that Kant's characterization of logic by its formality is his own innovation, not a holdover from tradition, and that his advocacy of this characterization dates from the beginning of his articulation of transcendental idealism. This historical fact cries out for a philosophical explanation. In this section, I will argue that the historical connection between transcendental idealism and logical hylomorphism is no coincidence, for there is a deep and intimate *philosophical* connection between the two doctrines. Transcendental idealism *implies*

**Kant's Thesis:** General (i.e., 1-formal) logic must also be formal (i.e., 3-formal). That is, "...the universal and necessary rules of thought in general can concern merely its *form* and not in any way its *matter*" (JL:12).

Thus, a transcendental idealist must hold that general logic is formal.

Although Kant never explicitly states the grounds for his commitment to Kant's Thesis, we can reconstruct an argument from four Kantian premises. The first premise is that

**(TS)** Thought is intelligible independently of its relation to sensibility.

Though Kant holds that *knowledge* or cognition of an object requires both thought and sensibility, he holds that the contributions of the two faculties can be distinguished (KrV:A52/B76). Thus, "...if no intuition could be given corresponding to the concept, the concept would still indeed be a thought ..." (B146)—though in this case one would have merely "empty" concepts, "mere forms of thought, without objective reality" (B148). As Parsons 1969 points out, Kant's metaphysics of morals presupposes the possibility of such a "problematic" extension of thought beyond the bounds of sense (117). Kant insists that

...*for thought* the categories are not limited by the conditions of our sensible intuition, but have an unlimited field. It is only the *knowledge* of that which we think, the determining of the object, that requires intuition. In the absence

of intuition, the thought of the object may still have its true and useful consequences, as regards the subject's *employment of reason*. The use of reason is not always directed to the determination of the object, that is, to knowledge, but also to the determination of the subject and of its volition—a use which cannot therefore be here dealt with. (B166 n.)

The second premise is that

**(CJ)** Concepts can be used only in judgment.

On Kant's view, "...the only use which the understanding can make of these concepts is to judge by means of them" (A68/B93).

The third premise is that

**(JO)** Judgment essentially involves the subsumption of an object or objects given in intuition under a concept.

Kant characterizes judgment as "...the mediate knowledge of an object, that is, the representation of a representation of it" (A68/B93). For Kant, what distinguishes a judgment (which is capable of being true or false) from a mere subjective association of representations (which is not) is that in a judgment, the representations are claimed to be "combined *in the object*" (KrV:§19). Thus on Kant's view there is no such thing as a judgment about concepts themselves, apart from their relation to an object.<sup>62</sup> The concepts in every judgment must relate finally to a "representation that is immediately related to an object" (A68/B93), that is, a singular representation of an object. And a singular representation of an object is just what Kant means by "intuition" (JL:§1).<sup>63</sup>

The fourth (and last) premise is that

**(OS)** Objects can be given to us only in sensibility. That is, for us (as opposed to God), all singular representations are sensible.

<sup>62</sup>Even the use of concepts in *analytic* judgments requires "relation to an object," although in this case we need not look beyond the concepts themselves to know the truth of the judgment and can therefore abstract from their relation to objects (A258/B314). That is, analytic judgments are still judgments about *objects*, not concepts (cf. Paton 1936:214 n. 3, Allison 1983:75).

<sup>63</sup>Kant sometimes adds that intuitions relate *immediately* to their objects (KrV:A320/B377). I take this to be a consequence of their singularity. For a discussion of this point with references to the literature, see Parsons 1969:112-114 and the Postscript.

“Our nature is so constituted,” Kant says, “that our *intuition* can never be other than sensible; that is, it contains only the mode in which we are affected by objects” (A51/B75; cf. A19/B33, A68/B92, A95, B146, A139/B178). In this we differ from God, whose intuition is “intellectual” or “original” (B72). God has singular representations of objects not through being affected by them, but by creating them.

From these four premises, we can obtain Kant’s Thesis. First, note that (CJ) and (JO) imply

**(CO)** A concept has content (objective validity, significance) only in so far as it applies to some object that could be given to us in intuition (that is, in a singular representation).

For according to (JO), the use of concepts in judgments consists in the subsumption under them of objects given in intuition. Hence, if no object could be given in intuition to which a concept might be applied, then that concept can have no use in judgment:

... the employment of a concept involves a function of judgment whereby an object is subsumed under the concept, and so involves at least the formal condition under which something can be given in intuition. If this condition of judgment (the schema) is lacking, all subsumption becomes impossible. For in that case nothing is given that could be subsumed under the concept. (KrV:A247/B304)

But since the *only* use concepts can have is their use in judgment (CJ), a concept for which no object could be given in intuition can have no use or objective significance at all (CO).

We demand in every concept, first, the logical form of a concept (of thought) in general, and secondly, the possibility of giving it an object to which it may be applied. In the absence of such object, it has no meaning and is completely lacking in content, though it may still contain the logical function which is required for making a concept out of any data that may be presented. (A239/B298; cf. A69/B93-4, B147, B148-9, A139/B178, A146/B185, A147/B186, A242/B299, A246/B302)<sup>64</sup>

Note that (CO) holds of mathematical concepts as well as empirical ones: mathematical concepts give one *a priori* cognition of objects, but only as regards their *forms* (as appear-

<sup>64</sup>The point is neatly summed up in a marginal note in Kant’s copy of the first *Critique* (at A19/B33): “The universal must be given in the particular. Through that it has significance” (Kant 1998:155).

ances); their content is contingent on “. . . the supposition that there are things which allow of being presented to us only in accordance with the form of that pure sensible intuition” (KrV:B147, cf. A239-40/B298-9).

From (CO) and (OS), it follows that

**(CS)** Concepts can have content only in relation to sensibility.

For concepts have content only insofar as they apply to objects that can be given in intuition, and intuition (for us) is sensible:

. . . the condition of the objective employment of all our concepts of understanding is merely the mode of our sensible intuition, by which objects are given us; if we abstract from these objects, the concepts have no relation to any object. (A286/B342)

. . . without the data of sensibility [the categories] would be merely subjective forms of the unity of understanding, having no object. (A287/B343)

From (TS), it follows that

**(LS)** General logic abstracts entirely from the relation of thought to sensibility. (A54/B78)

For if thought is intelligible independently of sensibility, then the norms governing thought as such cannot depend on thought’s relation to sensibility.

Finally, from (LS) and (CS) it follows that

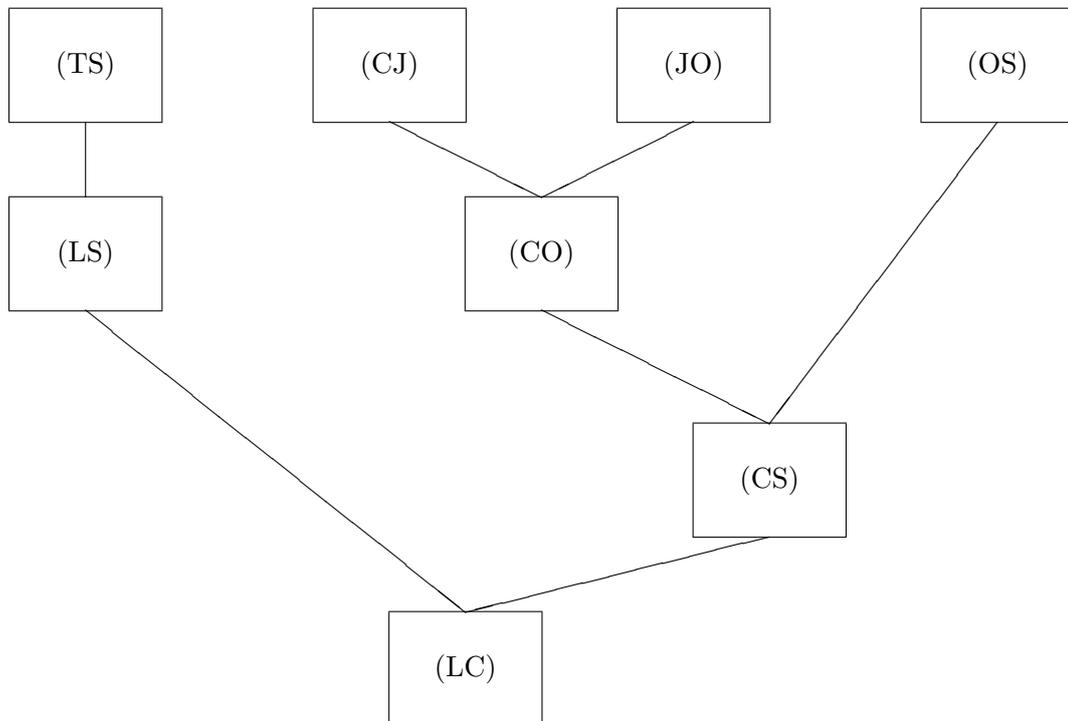
**(LC)** General logic abstracts entirely from the content of concepts.

As Kant would put it, general logic must be completely formal. In the terminology of chapter 3, a 1-formal logic must be 3-formal.

If this is the argument behind Kant’s inference from the generality of logic to its formality, we can explain why Kant does not draw this inference before 1773-5. By the dissertation of 1770, Kant has come to accept (TS) and (OS),<sup>65</sup> but he makes claims that are incompatible with (CO). Although he holds that the objects of the senses are “things *as they appear*,”

<sup>65</sup>For (OS), see §10, for (TS), §3.

Figure 4.2: The argument for Kant's Thesis.



he also claims that concepts can relate directly to “things *as they are*,” of which we can have no sensuous intuition (ID:§4), and hence no singular representation at all. But he gives no account of *how* concepts can relate to objects of which we can have no intuitions. It is dissatisfaction with *this* lacuna that starts Kant on the chain of reflections that leads eventually to transcendental idealism.<sup>66</sup> In 1772 he writes to Herz:

In my dissertation I was content to explain the nature of intellectual representations in a merely negative way, namely, to state that they were not modifications of the soul brought about by the object. However, I silently passed over the further question of how a representation that refers to an object without being in any way affected by it can be possible. I had said: The sensuous representations present things as they appear, the intellectual representations present them as they are. But by what means are these things given to us, if not by the way in which they affect us? And if such intellectual representations depend on our inner activity, whence comes the agreement that they are supposed to

<sup>66</sup>By “transcendental idealism,” I mean the doctrine that we can know objects only as they appear, not as they are in themselves.

have with objects—objects that are nevertheless not possibly produced thereby?  
(2/21/1772; Ak:X.129–35, trans. Kant 1967:72)

By 1775, Kant has resolved the difficulty by accepting (CO); he now explains the objectivity of pure concepts of the understanding through their applicability to *empirical* intuitions (as principles of order):

We have no intuitions except through the senses; *thus* no other concepts can inhabit the understanding except those which pertain to the disposition and order among these intuitions. (R:4673, trans. Guyer and Wood, Kant 1998:50, emphasis added)

Now Kant’s Thesis is inescapable, and Kant soon starts characterizing logic as “formal” (e.g., R:4676).

The upshot, I think, ought to be surprising to those who read Kant’s claims about the formality of logic as reiterations of a traditional view. Not only did Kant *invent* logical hylomorphism,<sup>67</sup> but he invented it *because his transcendental idealism required it!*

I do not make the converse claim—that logical hylomorphism requires transcendental idealism. I have not shown that one cannot take logic to be 3-formal unless one accepts all of Kant’s philosophy, or even that one cannot accept the Kantian argument (sketched above) unless one is a full-blown transcendental idealist. One might accept (TS), (CJ), (JO), and (OS) without taking objects to be “appearances” in Kant’s sense.<sup>68</sup> What I *have* shown is that logical hylomorphism is a substantive doctrine that stands in need of argument. If one is not willing to accept Kant’s argument from the generality of logic to its formality, one needs another argument for logical hylomorphism.

In calling logic “formal,” then, Kant is not giving a persuasive redefinition, but drawing a consequence from substantive philosophical premises and the neutral, accepted charac-

<sup>67</sup>Or reinvent it, if we want to count the medieval and ancient versions: see appendix A.

<sup>68</sup>I am grateful to Tyler Burge and Steve Engstrom for helping me see this. Note, however, that if one accepts (CS) without Kant’s “pure forms of intuition” and full transcendental idealism, one must give up at least one of these two Kantian claims: (i) mathematics has content, (ii) mathematics is non-empirical. Moreover, it is difficult to see what motivation one might have, apart from transcendental idealism, for accepting (TS) on top of (CS). John McDowell 1994 opposes (TS) precisely because of his commitment to (CS): “For a thought to be empty would be for there to be nothing that one thinks when thinks it, for it to lack what I am calling ‘representational content.’ That would be for it not really to be a thought at all . . .” (4).

terization of logic as “general.” What makes this hard to see, from our perspective, is that because of the huge influence of Kant’s philosophy, formality came to be seen as a *defining* characteristic of logic, even by philosophers who rejected transcendental idealism. “Formal logic” came to be seen as a specification of the *subject matter*, not a substantive claim about it. In the next section, we trace a bit of this story.

## 4.5 Kant as the source of modern logical hylomorphism

In arguing that logical hylomorphism is a Kantian innovation, I have looked mostly at Kant and his predecessors. I now want to offer further support for my thesis by looking at Kant’s *successors*. Nineteenth century discussions of logical hylomorphism commonly attribute the idea to Kant, and when they do not, they typically cite another source that does. The upshot is that Kant stands at the beginning of a long chain of historically connected uses of “formal” to characterize logic, a chain that runs right up to the present.

### 4.5.1 In Germany

In §186 of his *Wissenschaftslehre* (1837), Bolzano critically examines the idea that logic concerns the *form* of judgments, and not their matter—which he calls a doctrine of “the more recent logic.” The details of Bolzano’s criticisms do not concern us here. What is interesting is that nearly all of the explanations of the distinction between the form and matter of judgments he considers (and rejects) are from Kant (whom he considers first) or his followers (Metz, Hoffbauer, Kiesewetter, Jakob, Krug, Schaumann, Fries, Mehmel, van Calker).<sup>69</sup> Similarly, in §12, Bolzano considers the claim, made in “most contemporary treatises of logic,” that “in logic not the *content*, but the mere *form* of thought is to be

<sup>69</sup>Of the exceptions, only one is from a pre-Kantian writer (Baumgarten), whom Bolzano cites for the claim that the copula is the *partem formalem* of the judgment—hardly a counterexample to my claim that Kant is the originator of logical hylomorphism. The others are from Maass (a contemporary Wolffian critic of Kant) and Ulrich (a Wolffian who attempted “to reconcile and synthesize Kant’s critique with Leibniz’s metaphysics,” Beiser 1987:204), both of whom were directly influenced by Kant. Biographical data on these philosophers can be found in the *Allgemeine Deutsche Biographie*; see also Fries 1837:21 and Ueberweg 1882:51-3, who lists Jacob, Kiesewetter, Maass, Hoffbauer, Krug, and van Calker as logicians of the Kantian school.

investigated.” Again, all of the explanations he considers come from Kantians (Jakob, Hoffbauer, Metz, Krug). This evidence suggests that logical hylomorphism appeared to Bolzano (in 1837) as a newfangled doctrine, and one associated with Kant and his followers.<sup>70</sup>

We find much the same view at the end of the century, in Trendelenburg 1870 and Ueberweg 1882. Both thinkers associate the doctrine that logic concerns the mere form of thought, which they find pernicious, with Kant’s critical philosophy. Trendelenburg writes:

It is in Kant’s critical philosophy, in which the distinction of matter and form is thoroughly grasped, that *formal* logic is first sharply separated out; and properly speaking, it stands and falls with Kant. However, many who otherwise abandon Kant have, at least on the whole, retained formal logic. (1870:15, my translation)

Trendelenburg and Ueberweg seek to find room for an “Aristotelian” conception of logic, a middle road between the Kantian “subjective-formal logic,” which “sets the forms of thought outside of relation to the forms of being,” and the Hegelian “metaphysical logic,” which *identifies* the forms of thought and the forms of being (Ueberweg 1882:v, my translation; cf. Trendelenburg 1870:17). What is significant for our purposes is that the association of logical hylomorphism with Kantianism is still so strong at the end of the nineteenth century that these thinkers must take great pains to distinguish their version of the view from Kant’s.

#### 4.5.2 In Britain

The connection between Kantianism and logical hylomorphism is just as strong in Britain as it is in Germany, though not as self-conscious. In his influential lectures on logic, first composed and delivered in 1837-8 (and published posthumously in 1867), Sir William Hamilton articulates an essentially Kantian conception of logic, drawing heavily on German scholars

---

<sup>70</sup>Bolzano strenuously rejects views that restrict the domain of logic (which he takes to concern a Platonic realm of “propositions in themselves”) to *thought*: “. . . it is quite superfluous to speak of laws of thought when we could just as well deal with the conditions of truth itself” (§16). Nor will he accept that the propositions of logic lack content or do not contain truths, or that logic must abstract from differences between objects (§12). He is willing to grant, however, that logic is formal in the *schematic* sense: that is, it is concerned with *classes* of propositions that can be represented by proposition schemata (e.g., “Some A are B”) which can be called “forms” (§12, §186). In this sense, of course, the claim that logic is formal becomes trivial (§186).

influenced by Kant (primarily Esser and Krug).<sup>71</sup> Hamilton defines logic as “the Science of the Laws of Thought as Thought” (1867:3). He then glosses “Thought as Thought” as “the *form* of thought,” by which he means the *manner* in which an object is thought (15): “Now, when I said that Logic was conversant about thought considered merely as thought, I meant simply to say, that Logic is conversant with the form of thought, to the exclusion of the matter” (11). Hamilton justifies this limitation of logic to the form of thought with an argument borrowed from the German post-Kantians Esser and Krug.<sup>72</sup> Like Kant, Hamilton takes this conception of logic to rule out any use of logic as an *organon* or instrument of discovery (24, 32). As far as I can determine, none of these ideas appear in the work of British logicians before Hamilton’s 1833 review essay: Whately’s *Elements of Logic* (1826), the first major English logic text since Aldrich’s *Artis Logicae Compendium* (1691), is wholly innocent of logical hylomorphism and betrays no knowledge of the Kantian tradition in logic.<sup>73</sup> My claim is corroborated by Trendelenburg, who says that it was Hamilton who first brought Kantian formal logic to English soil (1870:15 n. 2), and by De Morgan, who remarks:

It is only of late years that, in this country at least, Kant’s definition has been clearly apprehended, and its truth sincerely felt. If the inquirer will look out

---

<sup>71</sup>Hamilton (a professor at Edinburgh) should not to be confused with his contemporary, the Irish mathematician Sir William Rowan Hamilton. The ideas elaborated in Hamilton’s lectures first appeared in an influential review article in the *Edinburgh Review* (Hamilton 1833).

<sup>72</sup>“The objects (the matter) of thought are infinite; no one science can embrace them all, and therefore, to suppose Logic conversant about the matter of thought in general, is to say that Logic is another name for the encyclopaedia—the *omne scibile*—of human knowledge. The absurdity of this supposition is apparent. But if it be impossible for Logic to treat of all the objects of thought, it cannot be supposed that it treats of any; for no reason can be given why it should limit its consideration to some, to the exclusion of others. . . . it follows that it must exclude from its domain the consideration of the matter of thought altogether; and as, apart from the matter of thought, there only remains the form, it follows that Logic, as a special science of thought, must be viewed as conversant exclusively about the form of thought” (16). The argument clearly derives ultimately from Kant’s argument that general logic can provide only a formal criterion of truth (JL:50-1). As it stands, it seems to depend on the unwarranted assumption that if logic concerned *all* objects of thought, it would have to yield *complete* knowledge of them. Nothing is said to rule out a third alternative: that logic tells us a few very general things about every object (about all logical “matter”).

<sup>73</sup>Whately does not invoke formality in his characterization of the science of logic in the Introduction. He does acknowledge the *schematic* formality of the syllogism (24, 14), but he is clear that schematic formality does not distinguish logic from arithmetic and algebra (14).

for English works preceding 1848, or thereabouts, which state Kant's definition as an existing thing, not to speak of adopting it, he will have some difficulty in finding one. (76 n. 1)

Interestingly, Hamilton seems not to recognize the Kantian provenance of his conception of logic, though he admits that Kant's "views of the nature and province of the science were peculiarly correct..." (1833:215). Indeed, he tries to present logical hylomorphism as a traditional view about logic. The Greek Aristotelians and Latin Schoolmen, he says, all delineated logic "... as a science of the form and not the matter of thought" (19-20). But instead of quoting logicians actually *saying* this, he argues that the medieval talk of logic as concerned with second intentions (in both its nominalist and realist varieties) "... merely varies and perplexes the expression, that the object of Logic is the formal laws of thought" (20).<sup>74</sup> He then claims (giving no evidence) that "[t]he same views, various in appearance, but, when analyzed, essentially the same, and essentially correct, may be traced through the Leibnitio-Wolfian school into the Kantian" (20). It is only his English predecessors that he finds to have misunderstood the "nature of logic" (21; cf. Hamilton 1833).

Hamilton's lectures were enormously influential: as Peter Heath writes,

... the fame of his teaching, and the infiltration of his pupils and followers into leading academic positions in Britain and America, had conferred on the Hamiltonian logic a prestige rivaled only by that of Mill, and reinforced by the appearance of his doctrines in such standard textbooks as Thomson's *Outline of the Necessary Laws of Thought*.... (De Morgan 1966:xii)<sup>75</sup>

Thus, in Thomson we find the claim that "*Pure logic is a science of the form, or of the formal laws of thinking, and not of the matter*" (1860:16-17):

<sup>74</sup>His editors provide one quote on his behalf: "Logicus solas considerat formas intentionum communes" (Albertus Magnus, *In De Anima*, I.i.7). But in this chapter Albert is talking about the distinction between logic and physics: his point is that logic, unlike physics, does not consider the matter of *things*, but just their common forms. The passage has nothing to do with a distinction between the form and matter of *thought*.

<sup>75</sup>In a 1952 letter, De Morgan calls Hamilton, with whom he is embroiled in a controversy over plagiarism charges, "the first logical name in Britain, it may be in the world" (letter to Spalding, quoted in De Morgan 1966:xii). Hamilton's fame is so great that De Morgan is delighted to be attacked by him "... in a printed book which will last" (letter to William Rowan Hamilton, quoted in De Morgan 1966:xviii).

The form is what the mind impresses upon its perceptions of things, which are the matter; form therefore means *mode of viewing* objects that are presented to the mind. When the attention is directed to any object, we do not see the object itself, but contemplate it in the light of our own prior conceptions. . . . The form then in this view is the mode of knowing; and the matter is the perception, or object we have to know. Hence, when we call Logic a science of the formal laws, or the form, of thinking, we mean that the science is only concerned with that which is essential to, and distinctive of, the thinking process. (21-3)

In the Preface to *The Laws of Thought*, Boole cites Thomson's book as one of the two best guides to the terms and "general object" of logic (1854:i). His own unpublished manuscripts are full of ruminations about the "formal" character of logic.<sup>76</sup> Similarly, in a manuscript logic text dating from 1860, De Morgan writes:

1. Logic analyses the *forms*, or *laws of action*, of thought.
2. Logic is *formal*, not *material*: it considers the law of action, apart from the matter acted on. It is not *psychological*, not *metaphysical*: it considers neither the mind in itself, nor the nature of things in itself; but the mind in relation to things, and things in relation to the mind. (1860:153, cf. 1858:75)

For the understanding of the word "form," he footnotes Thomson, placing himself at the end of a chain of uses reaching back to Kant, via Thomson, Hamilton, and Esser.

Unlike Hamilton himself, De Morgan sees clearly that the hylomorphic characterization of logic is a Kantian innovation:

To [Aristotle] we owe such perpetual indication of the distinction of form and matter that many, including some who should have known better, have assigned the *form of thought* to him as his definition of *logic*, giving him the word into the bargain. But the definition was never distinctly conceived in that character

<sup>76</sup>See especially Boole's unpublished manuscript "On the Foundations of the Mathematical Theory of Logic and on the Philosophical Interpretation of its Methods and Processes," in Boole 1997:104, esp. 68, 70-72, 98. Much of this material bears the stamp of the Kantian conception of logical formality passed on by Hamilton. For instance, Boole distinguishes between a formal and material element in every concept, "the former more immediately connected with sense or experience the latter more directly related to the faculties by which we hold converse with the scientific forms of truth" (68) and holds that "[t]he validity... of the process of *reasoning* depends not at all upon the pictorial element in our concepts whereby the images of sense are reproduced but only upon the formal laws and relations of those concepts" (70). Boole thinks it significant in this connection that we can reason validly using his symbolic rules even when some of the formulas arrived at in the course of this reasoning do not admit of any coherent interpretation (72-3; cf. "Logic," Boole 1997:144-8).

until the last century, when it was propounded by a philosopher whose earliest studies had been in mathematics, which he had taught in conjunction with logic for fifteen years before he gave himself up to the study of the pure reason. (1858:76)

Similarly, Mansel 1851 (who acknowledges Kant and Hamilton as major influences) writes:

For a period of seventy years, *reckoning from the first publication of Kant's Critique of Pure Reason*, Formal Logic, in itself and in its relations to Psychology, has been elaborated by numbers of eminent writers in Germany, from whose labours the English student has, as yet, derived hardly any benefit. . . . Few who are acquainted with the various logical systems of modern times will hesitate to give a decided preference over all others to *the formal view of the science, which from the days of Kant has gradually been advancing to perfection*. (ii, iv, emphasis added)<sup>77</sup>

The hylomorphic characterization reappears in logic texts up into the twentieth century. As it becomes entrenched and comes to feel more like a neutral and uncontroversial characterization of logic than a substantive and disputable thesis, it loses its specifically Kantian meaning. Its staying power is enhanced by the plasticity of the notion of formality, which can be reinterpreted as schematic formality or as generality (in the sense of either 1-formality or 2-formality).<sup>78</sup> Thus, Hyslop 1892 follows Hamilton in taking logic to be concerned with “*formal* laws of thought,” but glosses these as “the laws which are not only essential to it, but which are the same whatever the subject-matter involved in our reasoning” (12). Jevons 1870 glosses “the science of the necessary forms of thought” (4) by reference to the schematic formality and generality of logic. Frege uses the word “formal” in at least one passage (1885:95) to denote generality or 1-formality. And J. N. Keynes, in his widely-read textbook (fourth edition, 1906), notes that “[i]t is *usual* to say that logic is *formal* in so far as it is concerned merely with the form of thought, that is, with our

<sup>77</sup>At the beginning of chapter one, Mansel announces that logic “. . . will be treated in the following pages, in accordance principally with the views of Kant, as the Science of the Laws of Formal Thinking” (1). Mansel’s conception of logic differs from Kant’s, however, in being avowedly psychologistic (vi). His “necessary laws of thought” depend on the particular constitution of the human mind and are necessary only for *human* thought (74).

<sup>78</sup>Mill complains about Hamilton’s invocation of matter and form in logic: “It is a pity that the only terms that he can find to denote the distinction, are a pair of the obscurest and most confusing expressions in the whole range of metaphysics” (454).

manner of thinking irrespective of the particular objects about which we are thinking ...” (2, emphasis added), then glosses formality as generality and abstractness: “. . . we become more and more formal as we become more and more general; and logic may be said to be more abstract, more general, more formal, than any other science, except perhaps pure mathematics” (2-3). In this way, what was originally a daring proposal connected with a specific program in philosophy is gradually converted into a commonplace, and the word “formal” comes to have the range of different meanings we distinguished in chapters 2 and 3.<sup>79</sup>

## 4.6 Conclusion

I have argued that logical hylomorphism is a Kantian innovation. That logic must be formal is not part of Kant’s definition of logic, but a substantive claim supported by premises from his critical philosophy. I have given three kinds of argument for this claim:

- *Exegetical*: Several key texts show Kant *inferring* the formality of its logic from its generality.
- *Historical*: Kant’s predecessors did not characterize logic as distinctively formal. Kant himself only began to do so at around the same time as he became a transcendental idealist.
- *Philosophical*: There is a good argument that a general logic must be formal, using premises from Kant’s critical philosophy.

---

<sup>79</sup>This is not to say that the characterization of logic as formal is universally accepted. As we have already seen (footnote 78), Mill rejects it. F. H. Bradley’s *Principles of Logic* (1883) contains a penetrating discussion of the idea that logical reasoning is “formal.” Bradley shows a good understanding of the difference between the claim that logic is *schematically* formal and the claim that it is 3-formal (abstracts from all content): “You cannot conclude, because a male proves fertile with every known female, that he therefore supplies the principle of fertility. That would be quite absurd; and it is always absurd, when a result appears from a pair of elements, to argue, Because the *specialty* of the element on one side does not affect the general type of the result, the other element is the sole cause of this type. *For something common to all the different cases may exist and may work from its material side*, and hence some matter after all may belong to the essence of the formal activity” (1883:519-20, emphasis added).

I have also sketched the transmission of the doctrine from Kant to the German and British logico-philosophical traditions.

If my thesis is correct, then those who reject transcendental idealism should be wary of logical hylomorphism as well. This is not yet to say that logical hylomorphism cannot be *defended* apart from transcendental idealism: perhaps it can be. I have argued only that Kant's Thesis must be true if transcendental idealism is true, not that Kant's Thesis must be false if transcendental idealism is false. We should be wary of logical hylomorphism not because it has been shown to be false, but because it enjoys an undeserved default plausibility. If we wish to do without transcendental idealism, then we have a choice: find an entirely different *motivation* for adopting Kant's Thesis and the resulting conception of logic, or reject Kant's Thesis and acknowledge that general logic need not abstract entirely from content.

In chapter 5, I show how Frege pursues the second option.

## Chapter 5

# FREGE AND THE FORMALITY OF LOGIC

In chapter 4, I argued that Kant takes logic to be 3-formal because his transcendental idealism drives him to accept

**Kant's Thesis:** General (i.e., 1-formal) logic must also be formal (i.e., 3-formal). That is, "...the universal and necessary rules of thought in general can concern merely its *form* and not in any way its *matter*" (JL:12).

Formality is not, for Kant, part of the definition of logic; it is a substantive consequence of logic's generality, together with several other systematic claims.

In this chapter, I will examine Frege's reasons for rejecting Kant's Thesis, and with it the claim that logic is "formal" (3-formal). Like Kant, Frege demarcates logic by its "generality" or 1-formality. But unlike Kant, Frege takes logic to be "about the world" in much the same way as physics is (though naturally it concerns more abstract or general features of the world than physics). For Frege, logic *is* (*pace* Kant) a source of substantive knowledge, knowledge about objects.

Frege's rejection of Kant's Thesis is interesting for several reasons.

- First, Frege is not criticizing Kant from an entirely alien philosophical position. He takes his fundamental epistemological picture from Kant, altering it only as much as

is necessary to accommodate his central departure from the master: the view that arithmetic reduces to logic. Thus Frege's rejection of Kant's Thesis is likely to tell us more about the ultimate philosophical grounds of the thesis than, say, Mill's or Bradley's rejections of it would.

- Second, Frege rejects Kant's Thesis because he rejects the Kantian doctrine that motivates it: the doctrine that thought cannot have content without relation to sensible intuition.<sup>1</sup> Frege argues that the objects and concepts of arithmetic are given to us by reason alone, with no help from sensibility:

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 our reason and, as its nearest kin, utterly transparent to it. (FA:§105; cf. §89)

But if the objects of arithmetic are given directly to reason, then the norms constitutive of the use of reason—that is, the laws of logic—cannot abstract entirely from content.

- Third, as Frege develops his logicism, he becomes increasingly explicit in his rejection of the Kantian view that logic abstracts from all content. In his early work, he is still (to some degree) under the spell of the Kantian conception of logic; in his later work, he gives very clear reasons for rejecting it.
- Finally, the collapse of Frege's logicism (as a result of the paradoxes) leaves us with an interesting problem: to what extent are Frege's reasons for rejecting Kant's Thesis still available to us?

I begin (section 5.1) with an examination of Frege's early works (1879-1883). These works contain no systematic discussions of the nature of logic; however, passing remarks reveal that Frege has not yet freed himself from the Kantian view that logic concerns only the form of thought, in abstraction from its content. In section 5.2, I show how Frege's

---

<sup>1</sup>(CS) in section 4.4, above.

introduction of logical objects in the *Grundlagen* (1884) forces him to reevaluate the Kantian conception of logic: while he continues to characterize logic as 1-formal (thereby ensuring that he is talking about the same thing as Kant), he is increasingly explicit that logic is *not* 3-formal. He begins to reflect seriously on the nature of logic and offers an explanation of how logical laws can be normative for thought as such without abstracting entirely from semantic content—that is, how Kant’s Thesis can be false. In section 5.3, I argue that by the *generality* of logic, Frege means its 1-formality, not its 2-formality. Frege does *not* think that logic is general in the sense of being indifferent to the particular identities of objects. Finally, in section 5.4, I consider whether Frege’s reasons for rejecting Kant’s Thesis are vitiated by Russell’s paradox and Frege’s consequent abandonment of “logical objects.”

## 5.1 The status of logic in Frege’s early works

In the *Begriffsschrift* (BGS, 1879) and other early works, Frege seems not to be fully aware of the ways in which his conception of logic will have to diverge from Kant’s, and as a result there is very little explicit reflection on the nature of logic. “In the *Begriffsschrift*,” Sluga 1980 observes, “Frege had given almost no reasons why the formulas should be considered expressions of logical laws” (102). The reason for this lack of reflection, I suggest, is that in the *Begriffsschrift* Frege has yet to make a decisive break with the Kantian doctrine that objects can be given to us only in intuition. The logical laws presented in the *Begriffsschrift* do not allow for the introduction of “logical objects,” such as extensions. Nor has Frege yet begun to talk of the truth values or the senses of expressions as (nonsensible) objects. As a result, he has not yet been confronted with an outright *incompatibility* between the Kantian conception of logic and his own project. What we find in the early works is an unreflective wavering between a Kantian conception of logic (as both 1- and 3-formal) and something akin to Frege’s later conception (as 1-formal but not 3-formal).

Frege’s *primary* characterization of logic in the BGS is as 1-formal—as providing constitutive norms for thought as such, apart from any relation to intuition:

The firmest method of proof is obviously the purely logical one, which, disregarding the particular characteristics of things, is based solely upon the laws on which all knowledge rests. (BGS:Preface)

...I had first to test how far one could get in arithmetic by logical deductions alone, supported only by the laws of thought, which transcend all particulars. ...So that something intuitive could not squeeze in unnoticed here, it was most important to keep the chain of reasoning free of gaps. (BGS:Preface)<sup>2</sup>

In his later work, as we will see, Frege will argue that the 1-formality of logic does not entail its contentlessness, and at one point in the BGS he anticipates this view. He introduces his purely logical proof of several theorems in the theory of sequences as follows:

...we see in this example how pure thought (regardless of any content given through the senses or even given a priori through an intuition) is able, all by itself, to produce *from the content which arises from its own nature* judgments which at first glance seem to be possible only on the grounds of some intuition. We can compare this to condensation by which we succeed in changing air, *which appears to be nothing to the childlike mind*, into a visible drop-forming fluid. (BGS:§23, emphasis added)

But several passages in the earlier works indicate that Frege has not completely broken free from the Kantian conception of logic as abstracting entirely from the content of thought and dealing only with its form. In a quartet of articles devoted to explaining the Begriffsschrift and its purpose,<sup>3</sup> Frege makes extensive use of the form/content contrast. One of the main points in which the Begriffsschrift differs from Boole's calculus, Frege claims, is that while Boole's calculus is intended "... to present the logical form with no regard whatever for the content" (PW:47), the Begriffsschrift is intended as a tool for "the *expression of a content*" (PW:12; cf. SJ:86, ACN:90-1). In light of Frege's later work, it is tempting to read such claims as asserting (*pace* Kant) that logic alone suffices to "express" or "present" a content. But that is not Frege's point here. His point is rather that the Begriffsschrift can represent

<sup>2</sup>Propositions proved by means of pure logic, Frege emphasizes, have far greater *generality* than propositions proved by appeal to intuition: while the former apply to whatever can be thought, the latter are valid "only in the domain of the particular intuition upon which they were founded" (BGS:§23).

<sup>3</sup>SJ, ACN, and the unpublished manuscripts "Boole's logical Calculus and the Concept-script" (PW:9-46) and "Boole's logical Formula-language and my Concept-script" (PW:47-52). All four articles date from 1881-2.

the way in which more complex mathematical concepts are constructed from simpler ones: “... the content is not just indicated but is constructed *out of its constituents* by means of the same logical signs as are used in the computation” (PW:35, emphasis added). Frege’s whole discussion of the presentation of content presupposes that in addition to the logical signs, which reveal the “form” or structure of the complex content, there are signs indicating the simple non-logical contents that fill these forms. Every language, Frege claims, must have two components:

That is, we may distinguish the *formal part* which in verbal language comprises endings, prefixes, suffixes and auxiliary words, from the *material part* proper. The signs of arithmetic correspond to the latter. What we still lack is the *logical cement* that will bind these building stones firmly together. (PW:13, emphasis added)

While ordinary arithmetic contains only the *material* component, Boole’s calculus contains only the *formal* component. What the Begriffsschrift does, Frege maintains, is to “... supplement the signs of mathematics with a formal element” (PW:13, cf. PW:47), thereby encompassing *both* mathematical signs like “+” *and* logical signs like the conditional in a single language. In such a language, Frege says, “... the existing symbols [of mathematics] correspond to the word-stems of [ordinary] language; while the symbols I add to them are comparable to the suffices and [deductive] formwords [*Formwörter*] that logically interrelate the contents embedded in the stems” (ACN:93).<sup>4</sup> Thus, a proper “conceptual notation”

... must have simple modes of expression for the logical relations which, limited to the necessary, can be easily and surely mastered. These *forms* must be suitable for combining most intimately with a content. ... The *symbols for denoting content* are less essential. They can be easily created as required, once the general [logical] *forms* are available. (SJ:88, emphasis added).

Frege even goes as far as to describe the logical framework of the Begriffsschrift as “a

---

<sup>4</sup>The distinction is made even more pronounced by the peculiar symbolism of the Begriffsschrift. In Frege’s symbolism, there is a generic difference in appearance between logical and non-logical signs: logical signs (except identity and placeholders for generality) are represented by a two-dimensional structure of lines in which letters representing non-logical content can be embedded.

perspicuous representation of the *forms of thought*” (SJ:89, emphasis added).<sup>5</sup> In sum, although Frege acknowledges that the contents of some concepts have a sophisticated logical articulation, he never claims in these articles that a content can be presented using *only* formal or logical signs. He claims only that his Begriffsschrift “has a more far-reaching aim than Boolean logic, in that it strives to make it possible to present a content *when combined with arithmetical and geometrical signs*” (PW 46, emphasis added).

It is true that Frege is committed to according theorems of pure logic a kind of “content.” All such formulas follow the “content stroke,” and “*Whatever follows the content stroke must always have an assertible content*” (BGS:§2). But there are indications that Frege—no doubt under the spell of the Kantian characterization of logic—regards the content of theorems of pure logic as somehow degenerate, as not really content at all. Of his logical axioms, he says: “That my sentences have enough content, *in so far as you can talk of the content of sentences of pure logic at all*, follows from the fact that they were adequate to the task” (PW:38, emphasis added). The qualification here—and even perhaps the comparison to condensation at BGS:§23—reveals that Frege still has one leg on the Kantian side of the fence. Indeed, the way Frege defines “conceptual content” (*begrifflicher Inhalt*) in *Begriffsschrift* §3 implies that the “content” of theorems of pure logic is degenerate. Instead of defining conceptual content directly, Frege sets down a criterion for *sameness* of conceptual content: two judgments have the same conceptual content just in case “... [all] the consequences which can be derived from the first judgment combined with certain others can always be derived also from the second judgment combined with the same others” (BGS:§3). I think we must understand “can be derived” here as “can be derived using the laws of logic;” for without the laws of logic, it is not determinate what can be derived from what—at least when logical vocabulary is involved.<sup>6</sup> But then it seems to follow that all

<sup>5</sup>Sluga 1980 suggests that this language is “...clearly borrowed from Kant and used without systematic reflection” (108). Sluga points out many parallels between Frege’s justification for his Begriffsschrift in SJ and Trendelenburg 1867’s proposal for scaling down the Leibnizian project of a “universal characteristic” by restricting it to “formal side of thinking” (1980:49-52). There is little doubt that Frege was influenced by the Trendelenburg article, which he footnotes in the Preface of BGS.

<sup>6</sup>This is not to deny that there may be some “material” relations of derivability in place, e.g.

theorems of pure logic have the same content. For let  $A$  and  $B$  be any two theorems of pure logic. Let  $\Gamma$  be any set of “other” sentences. Suppose  $D$  can be derived from  $\{A\} \cap \Gamma$ . Then  $D$  can also be derived from  $\Gamma$  alone, since  $A$  can be derived using only the laws of logic. Hence  $D$  can be derived from  $\{B\} \cap \Gamma$ .<sup>7</sup> A symmetrical argument shows that anything derivable from  $\{B\} \cap \Gamma$  is derivable from  $\{A\} \cap \Gamma$ . Hence  $A$  and  $B$ —and likewise *all* theorems of pure logic—have the same “conceptual content.” Given this consequence, it would have been reasonable for Frege to hold that theorems of pure logic all have “null content”—a degenerate case of conceptual content.

It might be thought that the technical achievement of the BGS—the expression and justification of claims Kant would have regarded as based on pure intuition, such as the connectedness and transitivity of the relation of “following in a sequence,” using logical means alone—suffices to show that pure logic has content. But this conclusion would be too hasty. The technical results do not by themselves decide the issue of the 3-formality of logic. They do show that Kant was wrong about *something*, but they do not show *what* he was wrong about. There are two alternatives.

On the one hand, we might take Frege’s technical results to show that Kant was wrong to think that the fundamental concepts of arithmetic cannot be grasped without pure intuition. Some mathematical concepts *can* be grasped using purely logical means, and some important mathematical theorems can be justified without any appeal to intuition. Hence we must reject Kant’s general doctrine that thoughts without intuition are empty. We must also reject the Kantian conception of logic: if logic alone can give rise to contentful concepts, then it cannot be said to abstract entirely from the content of thought. This interpretation of the BGS’s technical results is natural in light of the later trajectory of Frege’s thought.

from “Bill is running” to “Bill’s legs are moving.” But these material relations of derivability will not settle whether, e.g., “ $\sim(A \ \& \ B)$ ” is derivable from “ $\sim B$ ”. (Here I am responding to a comment by Bob Brandom.)

<sup>7</sup>Here we have to assume that the relation of derivability is monotonic: i.e., that adding premises cannot spoil a good derivation. Brandom points out that since *material* derivability is not in general monotonic, this may not be a good assumption. On the other hand, Frege never talks about a non-monotonic relation of derivability, so it is not unreasonable to suppose that the one at issue in §3 is monotonic.

But it is not compulsory: we might still deny that Frege has succeeded in *expressing a content* using only the resources of pure logic. What the numbered propositions of the BGS express, we might claim, are mere *forms of thought*: more complex ones than Kant envisioned, but nonetheless mere forms or schemata. In order to express a content, we must add matter to these forms by instantiating their italic function and relation letters with particular contentful functions and relations. Until then, we have just played with representations; we have not secured their relation to objects, and hence we have not really said anything about the world.<sup>8</sup> It is true that Frege’s technical results undermine some of the mathematical *motivation* for Kant’s view that content requires relation to intuition: as Friedman has argued, Kant is driven to that view in part by his realization that Aristotelian logic does not suffice for the expression of mathematical concepts and the justification of mathematical theorems.<sup>9</sup> But one might retain the view that content requires relation to sensible intuition for *philosophical* reasons, even without this mathematical motivation. Frege’s technical successes in expressing mathematical notions with purely logical means do not by themselves *demand* a revamping of the Kantian picture (and its associated conception of logic), though they do provide an intelligible *motivation* for such a revamping.<sup>10</sup>

<sup>8</sup>Frege’s substitutional account of generality (BGS:§11) makes this line possible. The judgment “ $\forall x\Phi x$ ” is for Frege “*the judgment that the function  $[\Phi]$  is a fact whatever we may take as its argument,*” where “function” and “argument” are syntactic entities (§9). Frege’s account of generality over functions follows this pattern. Thus, to assert “ $\forall F\forall xFx$ ” is to assert that every instance of the schema “ $Fx$ ” holds. It is therefore open to Frege to take his propositions in the general theory of sequences to have much the same status as, say, a schematic presentation of the principle of excluded middle (“nothing is both F and not-F”) would have for Kant. That is: they do not say anything by themselves; only their instances have content. Note that this Kantian line is *not* compatible with Frege’s later account of the quantifiers as second-level functions: on this later account, to assert “ $\forall F\forall xFx$ ” is to assert something about the totality of first-level concepts, construed objectively as functions from objects to truth values.

<sup>9</sup>See Friedman 1992:ch. 1–2. Friedman argues that the expressive capacity of Frege’s iterated quantifiers makes it possible to *describe* structures which for Kant could only be represented by *exemplifying* them in intuitive construction. See section 5.2.1, below.

<sup>10</sup>Indeed, it was not uncommon for mathematicians engaged in projects of rigorization to describe the results of their work as mere *forms* or *schemata* of concepts. For example, in his *Ausdehnungslehre* (1844)—probably the first *explicitly* abstract work in geometry—Grassmann defines “structures” of “elements” and “operations” without reference to anything spatial (see Nagel 1979:218). He describes his new science of geometry as a “formal” science, not a “real” one—in fact, “the general science of pure forms” (see Nagel 1979:216). Thus, instead of claiming that Kant was wrong to associate content with relation to intuition, Grassmann accepts the Kantian framework and re-describes his work as the articulation of contentless forms. Similarly, in a letter to Frege

The passages discussed above suggest that as late as 1882, Frege is ambivalent between these two interpretations of the BGS as a response to Kant. He leans toward the first, but some of his remarks suggest the second. I suggest that he has not yet seen the extent to which he will be driven to depart from Kant.

In sum, I have argued that

- The reason that Frege does not provide an explicit discussion of the nature of logic in the early works is that he has not yet seen clearly the need to articulate a conception of logic that diverges from Kant's.
- Several passages reveal that he is still under the spell of the Kantian conception of logic as 3-formal.
- However, other passages show that Frege is beginning to reject the Kantian conception.
- The technical results of the BGS do not by themselves decide the issue.

In the next section, I will argue that when Frege begins to grapple explicitly with the issue starting in 1884, he comes down decisively against the Kantian conception of logic.

## 5.2 Frege's rejection of Kant's Thesis

### 5.2.1 Frege and Kant on logic and arithmetic

The technical project outlined in Frege's *Foundations of Arithmetic* (FA, 1884) is considerably more ambitious than that carried out in the *Begriffsschrift*. Frege now aims to show not just that (generalizations of) theorems Kant would have regarded as arithmetical can be proved in pure logic, but that arithmetic has its own *objects* (the numbers), which are given to us by logical means alone. This project—unlike that of the BGS—*demand*s a

---

dated December 29, 1899, Hilbert describes his *Grundlagen der Geometrie* as providing “. . . only a scaffolding (schema) of concepts together with their necessary connections,” which can be “filled” in various ways by interpreting the primitives: “e.g., instead of points, think of a system of love, law, chimney-sweep. . . which satisfies all axioms; then Pythagoras' theorem also applies to these things” (C:42).

break with the Kantian doctrine that thoughts have content only through their relation to sensible intuition, and Frege acknowledges his disagreement with Kant on this issue:

I must also protest against the generality of Kant's dictum: without sensibility no object would be given to us. Nought and one are objects which cannot be given to us in sensation. (FA:§89)

Yet he continues to defer to Kant for his basic epistemological framework:<sup>11</sup>

I have no wish to incur the reproach of picking petty quarrels with a genius to whom we must all look up with grateful awe; I feel bound, therefore, to call attention also to the extent of my agreement with him, which far exceeds any disagreement. To touch only upon what is immediately relevant, I consider Kant did great service in drawing the distinction between synthetic and analytic judgments. In calling the truths of geometry synthetic and *a priori*, he revealed their true nature. And this is still worth repeating, since even to-day it is often not recognized. If Kant was wrong about arithmetic, that does not seriously detract, in my opinion, from the value of his work. His point was, that there are such things as synthetic judgments *a priori*; whether they are to be found in geometry only, or in arithmetic as well, is of less importance. (FA:§89)

Thus Frege accepts Kant's view of the options for mathematical knowledge (analytic *a priori*, synthetic *a priori*, synthetic *a posteriori*—FA:§3), and he accepts Kant's view that geometry is synthetic *a priori*. The disagreement, which Frege portrays as an intramural one, concerns only the status of arithmetic. While Kant takes arithmetic to be synthetic *a priori*, Frege thinks that it rests on general logic and definitions alone and is therefore analytic. Frege sums the results of FA this way: "From all the preceding it thus emerged as a very probable conclusion that the truths of arithmetic are analytic and *a priori*; and we achieved an improvement on the view of Kant" (§109).

But clearly the issue has not been joined unless Frege and Kant can agree on what counts as logic. And it is not obvious that they can. As Friedman 1992 has emphasized, Frege's logic has considerably more expressive power than Kant's. Consider the idea that "for every  $n$  there is a number  $n+1$ " (126), which can be expressed in a Fregean logic as

<sup>11</sup>Kitcher 1979, Sluga 1980:43-4, and Weiner 1990 (esp. ch. 2) have emphasized the extent to which Frege's epistemological project is embedded in a Kantian framework. This is not to deny that Frege's project is also motivated by issues in contemporary mathematics: see especially Tappenden 1995.

(I)  $\forall x \exists y (y = x + 1)$ .

Kant's logic—Aristotelian term logic with a simple theory of disjunctive and hypothetical propositions added on—has no resources for representing this kind of quantifier dependence. Hence for Kant, Friedman argues,

The only way even to think or represent this proposition—so as, in particular, to engage in rigorous arithmetical reasoning thereby—is by means of our possession of the successor function itself: in Kant's terms, by our capacity successively to iterate any given operation. This, for Kant, presupposes the pure intuition of time . . . . (126)

We are inclined to say that Kant's position rests on his *ignorance* of the true logic. But it is not obvious that teaching Kant the Begriffsschrift would have changed his mind. It would have been open to him to claim that Frege's "logical" system is merely another (intuitively based) mathematical system, and that the meaning of the iterated quantifiers can only be grasped through construction in pure intuition.<sup>12</sup> As Dummett observes, "It is . . . not enough for Frege to show arithmetic to be constructible from some arbitrary formal theory: he has to show that theory to be logical in character, and to be a correct theory of logic" (1981:15). Kant might have argued that Frege's expansion of logic was just a change of subject, just as Poincaré 1908 charged that Russell's "logical" principles were really intuitive, synthetic judgments in disguise:

We see how much richer the new logic is than the classical logic; the symbols are multiplied and allow of varied combinations which are no longer limited in number. Has one the right to give this extension to the meaning of the word *logic*? It would be useless to examine this question and to seek with Russell a mere quarrel about words. Grant him what he demands, but be not astonished if certain verities declared irreducible to logic in the old sense of the word find themselves now reducible to logic in the new sense—something very different.

We regard them as intuitive when we meet them more or less explicitly enunciated in mathematical treatises; have they changed character because the meaning of the word logic has been enlarged and we now find them in a book entitled

---

<sup>12</sup>This line is not so implausible as it may sound. For consider how Frege explains the meaning of the (iterable) quantifiers in the BGS by appealing to the substitution of a potentially infinite number of expressions into a linguistic frame (Frege 1879).

*Treatise on Logic?* (1908:ch. 4, §11, 461).<sup>13</sup>

Kant would have a strong theoretical basis for denying that Frege's Begriffsschrift is a system of *logic*: his claim that logic must be *formal*. He could regard Frege's demonstration (or near-demonstration) that the substantive science of arithmetic can be obtained from the Begriffsschrift and definitions alone as a proof that the Begriffsschrift is not formal—and hence not a logic. As Hao Wang 1957 remarks, Frege's reduction “cuts both ways”: “. . . if one believes firmly in the irreducibility of arithmetic to logic, he will conclude from Frege's or Dedekind's successful reduction that what they take to be logic contains a good deal that lies outside the domain of logic” (80).

The dialectic would be irresolvable if formality were part of Kant's *definition* of logic. Then there would be no way to exonerate Frege from the charge of “changing the subject,” and no way to make sense of his claim to have overturned Kant's view on the status of arithmetic. But as I argued in chapter 4, Kant regards his claim that logic is formal as a substantive one, not a matter of meanings. Kant takes logic to be formal because he has to, *given* his transcendental idealism. This suggests a way to blunt one edge of Wang's double-edged sword. In the next section, I will argue that Frege demarcates logic in the same way as Kant does, by appealing to its “generality” or 1-formality. Thus Frege can reject Kant's claim that logic is formal without “changing the subject,” provided that he rejects enough of Kant's more general philosophical theses to block the argument for Kant's Thesis.

---

<sup>13</sup>Poincaré uses this point to argue that the logicians are wrong to claim that they have “ruined the Kantian theory of mathematics” (471). On Poincaré's criticisms of logicism, see Parsons 1965 and Goldfarb 1988. Boolos 1985 considers a similar “Kantian” charge that Frege's “logical” moves in fact presuppose an appeal to intuition (see esp. 161-3). In the same vein, Hintikka 1965 suggests that only *some* first-order quantificational inferences should count as analytic in Kant's sense: those that do not require the use of singular representations not contained in the premises (i.e., the “arbitrary objects” introduced by existential instantiation). And Quine has famously charged that second-order logic is really “set theory in sheep's clothing” (1986:66).

### 5.2.2 Frege's characterization of logic as 1-formal

Where it is important to distinguish the logical from the non-logical, Frege consistently invokes the idea that logical laws apply to anything *thinkable* and are therefore normative for thought *as such*. This characterization of the logical comes out most clearly in the passages in which Frege argues that arithmetic, because it applies to everything that can be thought (regardless of whether it is intuitable), must have a logical basis. In FA:§14, for example, Frege contrasts arithmetic with geometry along these lines:

For purposes of conceptual thought we can always assume the contrary of some one or other of the geometrical axioms, without involving ourselves in any self-contradictions when we proceed to our deductions, despite the conflict between our assumptions and our intuition. The fact that this is possible shows that the axioms of geometry are independent of one another and of the primitive laws of logic, and consequently are synthetic. Can the same be said of the fundamental propositions of the science of number? Here, we have only to try denying any one of them, and complete confusion ensues. Even to think at all seems no longer possible. The basis of arithmetic lies deeper, it seems than that of any of the empirical sciences, and even than that of geometry. The truths of arithmetic govern all that is numerable. *This is the widest domain of all; for to it belongs not only the actual, not only the intuitable, but everything thinkable. Should not the laws of number, then, be connected very intimately with the laws of thought?* (FA:§14, emphasis added)

By “laws of thought,” Frege does not mean the laws according to which our thinking does (or even can) proceed, but the laws according to which it *ought* to proceed (1918:58, PW:145). To say that the laws of logic are “laws of thought,” for Frege, is to say that “they are the most general laws, which prescribe universally the way in which one ought to think if one is to think at all” (1893:xv). So in connecting arithmetic with the laws of thought, Frege is saying that they are normative for thought as such, or 1-formal. When we entertain the negations of geometrical truths, we are no longer thinking correctly *about space*, but our thought cannot be faulted *qua* thought. When we entertain the negations of arithmetical truths, however, there is no longer any respect in which our thought is correct. There is no saying: we are no longer thinking correctly *about numbers*, but our thought cannot be

faulted *qua* thought.<sup>14</sup>

In his 1885 article “On Formal Theories of Arithmetic” (FTA), Frege again characterizes arithmetic and logic as 1-formal, and this time he uses the *word* “formal.” Arithmetic is a “formal theory,” he argues, not in the sense that it deals with “empty signs,” but in the sense that its theorems “...can be derived from definitions alone using purely logical means...” (FTA:94). Frege thinks that this claim is made plausible by arithmetic’s universally applicability:

As a matter of fact, we can count just about everything that can be an object of thought: the ideal as well as the real, concepts as well as objects, temporal as well as spatial entities, events as well as bodies, methods as well as theorems; even numbers in turn can be counted. . . . From this we may undoubtedly gather at least this much, that the basic propositions on which arithmetic is based *cannot apply merely to a limited area* whose peculiarities they express in the way in which the axioms of geometry express the peculiarities of what is spatial; *rather, these basic propositions must extend to everything that can be thought. And surely we are justified in ascribing such extremely general propositions to logic.*

I shall now deduce several conclusions from this *logical or formal* nature of arithmetic. (94-5, emphasis added)

By “formal” here, Frege appears to mean *1-formal*: logic and arithmetic are applicable to anything thinkable because they are normative for thought *as such*. Thus, although Frege will deny that logic is formal in the sense of abstracting entirely from the content of thought (3-formal), he can agree with Kant that logic is “general” or 1-formal. This agreement provides a basis for Frege’s *disagreement* with Kant about the reducibility of arithmetic to logic.<sup>15</sup>

---

<sup>14</sup>See also Frege’s letter to Anton Marty: “The field of geometry is the field of possible spatial intuition; arithmetic recognizes no such limitation. . . . Thus the area of the enumerable is as wide as that of conceptual thought, and a source of knowledge more restricted in scope, like spatial intuition or sense perception, would not suffice to guarantee the general validity of arithmetical propositions. And to enable one to rely on intuition for support, it does not help at all to let something spatial represent something non-spatial in enumeration; for one would have to justify the admissibility of such a representation.” (8/29/1882, 1980:100).

<sup>15</sup>Thus also Friedman 1988: “The principles and theorems of the *Begriffsschrift* are implicit in the requirements of any coherent thinking about anything at all, and this is how Frege’s construction of arithmetic within the *Begriffsschrift* is to provide an answer to Kant: arithmetic is in no sense dependent on our spatiotemporal intuition but is built in to the most general conditions of thought

### 5.2.3 “The legend of the sterility of pure logic”

The characterization of logic as 1-formal can be found even in the *Begriffsschrift* (see section 5.1, above). What is new (beginning in 1884) is an increasingly explicit rejection of the Kantian notion that logic is 3-formal.

This rejection begins with an attack on the Kantian view that logic is sterile (incapable of extending knowledge). After presenting the view that arithmetic is analytic, Frege notes:

But this view, too, has its difficulties. Can the great tree of the science of number as we know it, towering, spreading, and still continually growing, have its roots in bare identities? And how do the empty forms of logic come to disgorge so rich a content? (FA:§16)

Similarly, in the 1885 article, he writes:

If this formal theory is correct, then logic cannot be as barren as it may appear upon superficial examination—an appearance for which logicians themselves must be assigned part of the blame. (FTA:95)

He suggests that the source of the erroneous view that logic is barren is the fact that mathematical calculation can be done using mechanical rules, even by someone who does not understand the symbols. But this *syntactic* formality, Frege points out, does not imply 3-formality:

... it is possible for a mathematician to perform quite lengthy calculations without understanding by his symbols anything intuitable, or with which we could be sensibly acquainted. And that does not mean that the symbols have no sense; we still distinguish between the symbols themselves and their content, even though it may be that the content can only be grasped by their aid. (FA:§16)

Indeed, Frege argues, even if we could not define the arithmetical primitives in purely logical terms, and had to proceed in arithmetic as in geometry, starting from axioms with non-logical primitives—even then, the “prodigious development of arithmetical studies” should “... suffice to put an end to the widespread contempt for analytic judgements and to the legend of the sterility of pure logic” (FA:§17). For certainly it is an extension of itself. This, in the end, is the force of Frege’s claim to have established the analyticity of arithmetic” (84).

our knowledge to discover that all the theorems of arithmetic follow logically from a few basic starting axioms—or equivalently, that all the conditionals whose antecedents are the axioms and whose consequents are the theorems are true.

In FA:§88, Frege suggests that Kant’s view that logic cannot extend knowledge has its source in the poverty of Kant’s logic. Because his logic is limited to relations between subject and predicate and contains no provision for nested quantification, Kant’s paradigm definition of a concept is a list of characteristics. But such definitions are the “least fruitful” kind: if we think of the simple characteristics as the smallest regions on a map, all we can do with a list of characteristics is “. . . to use the lines already given in a new way for the purpose of demarcating an area.” Thus it was not unreasonable for Kant to deny that a logic limited to drawing out the consequences of such definitions—a process in which “[n]othing essentially new. . . emerges”—can extend knowledge. The case is very different, however, for Frege’s logic, which (through the use of nested quantifiers) allows us to draw “. . . boundary lines that were not previously given at all.” What we can infer from a genuinely “fruitful” definition, as opposed to a list of characteristics, “. . . cannot be inspected in advance; here, we are not simply taking out of the box again what we have just put into it. The conclusions we draw from it extend our knowledge. . . .”

Interestingly, none of the considerations Frege adduces in these passages (§17, §88) depend on the the FA’s doctrine that arithmetic (and so also logic) has its own objects. Thus Frege’s attack on the “legend of the sterility of pure logic” is independent of the success of the logicist reduction. But *immediately after* this discussion of fruitful definitions and the capacity of logic to extend knowledge, Frege makes explicit his rejection of “Kant’s dictum” that “without sensibility no object would be given to us” (FA:§89). It is likely that Frege’s break with Kant on this point provided the *impetus* or *motivation* for his reevaluation of the doctrine that logic is 3-formal, even if it is not essential to his *justification* for rejecting this doctrine.

### 5.2.4 “Logical form” in the FA

Frege does claim in the FA that “[w]hat is of concern to logic is not the special content of any particular relation, but only the logical form” (§70). Doesn’t this amount to a commitment to the 3-formality of logic? I will argue that it does not.

In order to carry out his project of defining number in purely logical terms, Frege must show that the concept of *equinumerosity*<sup>16</sup> or one-to-one correlation, a key component of the definition, can itself be defined in purely logical terms. Accordingly, he defines “the concept F is equinumerous to the concept G” as follows:

there exists a relation  $\phi$  such that

1. If  $d$  stands in the relation  $\phi$  to  $a$ , and if  $d$  stands in the relation  $\phi$  to  $e$ , then generally, whatever  $d$ ,  $a$ , and  $e$  may be,  $a$  as the same as  $e$ ,
2. If  $d$  stands in the relation  $\phi$  to  $a$ , and if  $b$  stands in the relation  $\phi$  to  $a$ , then generally, whatever  $d$ ,  $b$ , and  $a$  may be,  $d$  as the same as  $b$ . (§72)

But what makes this definition a definition in *purely logical terms*? Frege argues that

The doctrine of relation-concepts is thus, like that of simple concepts, a part of pure logic. *What is of concern to logic is not the special content of any particular relation, but only the logical form.*<sup>17</sup> And whatever can be asserted of this, is true analytically and known a priori. . . . we can take “ $a$  stands in the relation  $\phi$  to  $b$ ” as the general form of a judgement-content which deals with an object  $a$  and an object  $b$ . (FA:§70, emphasis added)

In particular, since we have abstracted from the content of the relation  $\phi$  in the definition of equinumerosity (by generalizing over it and making it “indefinite”), it need not be conceived as a spatio-temporal relation (FA:§80).

Does the claim that logic is concerned with “the logical form” of a relation, not its “special content,” amount to an endorsement of the view that logic is 3-formal? No. Frege does *not* say that logic abstracts from the content of relations, but rather that logic abstracts from the *special* (*besondere*) content of relations. He thereby implies that logic *does* concern itself with the *general* content of relations: something Kant would never say.

<sup>16</sup> *Gleichzähligkeit*, translated “equality” by Austin.

<sup>17</sup> “Es kommt hier nicht der besondere Inhalt der Beziehung in Betracht, sondern allein die logische Form.”

### 5.2.5 Frege's rejection of 3-formality

For clarification on this issue, we must turn to Frege's 1906 polemic against Korselt and Hilbert on the foundations of geometry (FG2). In his monumental *Grundlagen der Geometrie* (1899), Hilbert had presented a complete axiomatization of Euclidean geometry and proved the independence of each axiom from the rest by giving non-geometrical interpretations of the geometrical primitives ("point," "line," "between," etc.) that made one axiom false and the others true. (Similar methods had long been used to establish the consistency of the various systems of non-Euclidean geometry.) For reasons we need not go into here, Frege considers the Hilbertian approach conceptually muddled. He thinks that Hilbert has proved only the independence of certain second level concepts, not the independence of the real Euclidean propositions (FG2:402). Yet Frege acknowledges the mathematical interest of independence proofs, and in the third part of his paper, he tries to develop a logically hygienic approach to proving that one thought is independent of a group of thoughts—that is, cannot be inferred from them using logical laws (FG2:423-4).

Frege's first observation is that this investigation cannot be carried out using the resources of the *Begriffsschrift* alone: some new "basic laws" will be required. For the independence results concern *thoughts* (the senses of sentences, not the sentences themselves), and mathematics is not *about* thoughts (426). The *Begriffsschrift*, a tool for doing mathematics, contains no primitive symbols that refer to thoughts and no laws governing relations between them (cf. PW:122). At the risk of anachronism, we might describe Frege's point as the realization that metalogic is distinct from logic. Contrast Hilbert's approach: his independence proofs are just more mathematics; there is no explicit codification of the semantic principles used.<sup>18</sup>

What new basic laws are required for a rigorous scientific treatment of logical independence? First, we will need laws relating logical dependence to truth: for instance, "If the thought G follows from the thoughts A, B, C by a logical inference, then G is true"

<sup>18</sup>I owe this point to Ricketts 1997:181. See also Tappenden 1997 for a rather different (though not wholly incompatible) perspective on the same issues.

(FG2:426).<sup>19</sup> But these will not suffice:

We need yet another law which is not expressed quite so easily. Since a final settlement of the question is not possible here, I shall abstain from a precise formulation of this law and merely attempt to give an approximation of what I have in mind. One might call it an emanation<sup>20</sup> of *the formal nature of logical laws*. (FG2:426, emphasis added)

The law in question lays down conditions for the preservation of logical dependence under translation. Frege aims to capture the idea (central to the use of alternative “interpretations” of geometrical primitives to prove independence) that logical dependence is insensitive to the non-logical content of thoughts. What the “formal nature of the laws of logic” suggests is that logic abstracts from *all* content: “. . . as far as logic itself is concerned, each object is as good as any other, and each concept of the first level as good as any other and can be replaced by it, etc.” (FG2:427-8). If this were the case, it would be easy to formulate the translation invariance criterion for logical dependence: logical dependence must be preserved through *any* translation that preserves the semantic categories of the contentful vocabulary and the truth of the premises.

But Frege rejects this criterion, on the grounds that “logic is not as unrestrictedly formal as is here presupposed,” i.e., not 3-formal:

If it were, then it would be without content. Just as the concept *point* belongs to geometry, so logic, too, has its own concepts and relations; and it is only in virtue of this that it can have a content. *Toward what is thus proper to it, its relation is not at all formal*. No science is completely formal; but even gravitational mechanics is formal to a certain degree, in so far as optical and chemical properties are all the same to it. . . . To logic, for example, there belong the following: negation, identity, subsumption, subordination of concepts. And here logic brooks no replacement. It is true that in an inference we can replace Charlemagne by Sahara, and the concept *king* by the concept *desert*, in so far as this does not alter the truth of the premises. But one may not thus replace the relation of identity by the lying of a point in a plane. Because for identity there hold certain logical laws which as such need not be numbered

<sup>19</sup>Frege’s notion of independence has the peculiarity that *false* thoughts are independent of all others (FG2:424). This is because Frege holds that “[o]nly true thoughts can be premises of inferences” (FG2:425). He understands “deductions from a hypothesis” not as inferences proper, but as assertions of conditionals.

<sup>20</sup>*Ausfluß*. Ricketts translates “upshot” (1997:182).

among the premises, and to these nothing would correspond on the other side. Consequently a lacuna might arise at that place in the proof. One can express it metaphorically like this: About what is foreign to it, logic knows only what occurs in the premises; about what is proper to it, it knows all. Therefore in order to be sure that in our translation, to a correct inference on the left there again corresponds a correct inference on the right, we must make certain that in the vocabulary to words and expressions that might occur on the left and *whose references belong to logic*, identical ones are opposed on the right. (FG2:428, emphasis added)

The fact that logic “has its own concepts and relations” (negation, identity, subordination of concepts, etc.)—that is, concepts and relations to which its laws are *not* indifferent—shows, in Frege’s view, that logic does not abstract entirely from content.<sup>21</sup>

Indeed, Frege holds that “[n]o science is completely formal.” Instead, there is a *partial ordering* of formality, depending on the concepts and relations to which the laws of each science are indifferent. Where A is a science, let I(A) be the set of concepts and relations to which the laws of A are indifferent. Then A is *more formal* than B (in the sense in which Frege is using the word here) just in case I(B) is a proper subset of I(A). This relation gives rise to a downward-branching tree structure, with logic at the top (since logical concepts and relations are employed in every science).<sup>22</sup> Thus, logic is the *most formal* science, in the sense that it abstracts from more concepts and relations than any other science, but it is not *completely formal*, because it does not abstract from *all* concepts and relations. In order to state the translation invariance criterion for logical dependence, then, a precise demarcation of the *logical* concepts and relations is needed: “. . . it will have to be determined what counts as a logical inference and what is proper to logic” (FG2:429).<sup>23</sup> Not knowing how to give such a demarcation, Frege abandons his metalogical investigations at this point

<sup>21</sup>We must take “concepts and relations” here to include second level concepts and relations, like subordination, as well as first level ones, like identity.

<sup>22</sup>To be more precise, truth-functional logic will be at the top, with quantificational logic beneath to it, and various extensions of these (e.g., modal, tense, identity) branching off from these.

<sup>23</sup>As Ricketts points out, nothing like this is needed for Frege’s logicist program (1997:184). All Frege needs to claim for that purpose is that the fundamental notions of the Begriffsschrift *are* logical notions; he does not need to claim that they are the *only* logical notions, and he does not need to say what it is to be a logical notion, except in a very general way (e.g., by invoking 1-formality or universal applicability).

(FG2:429).<sup>24</sup>

The crux of Frege’s disagreement with Kant about the 3-formality of logic is evidently his claim that logic has its own concepts and relations. Kant would agree, of course, that logical dependence is not preserved by translations unless the logical constants are kept fixed. But for Kant, the logical constants indicate *functions* or *forms* of judgment—ways in which representations are united in thought (in the transcendental unity of apperception) to form judgments. For example, in the judgment that all humans are mortal, the “all” says something about how the concepts *human* and *mortal* are combined in the judgment. From this perspective, there is no temptation to say that logic has its own proprietary concepts, like negation and concept subordination. These are not concepts at all, but ways in which concepts can be put together in thought. The view of the “formal nature of the laws of logic” which Frege rejects—the view that “. . . as far as logic itself is concerned, each object is as good as any other, and each concept of the first level as good as any other and can be replaced by it, etc.” (FG2:427-8)—has its proper place in this Kantian framework, according to which the logical vocabulary of a sentence describes a contentless *form* or framework, so that *all* content resides in the concepts that are unified by this framework to form a single judgment.

For Frege, by contrast, the logical constants refer to genuine concepts and relations. In the long FG2 passage quoted above, Frege emphasizes this point by saying that the *references* of the logical constants “belong to logic.” For example, just as “humans” in “all humans are mortal” refers to a concept that is true of every human, so “all” refers to a second level concept which is true of every concept that applies to every object. Because concepts are conceived as *functions* (see FC), there is no need to invoke “functions of thought” to explain how *Socrates* and *mortal* can be “stuck together” to form the thought *that Socrates is mortal*: the concept *is mortal* is nothing but a function from objects (including Socrates) to truth values; to speak figuratively, it contains in itself the possibility of its application to

---

<sup>24</sup>He never resumes it. In a 1910 letter to Jourdain, he writes: “The unprovability of the parallel axiom cannot be proved,” citing FG2 (quoted in Ricketts 1997:185).

Socrates. If there is anything corresponding to the Kantian “functions of thought” in Frege’s picture, then, it is the operations of functional application and its converse, functional abstraction. What would be left if we abstracted from *all* contentful function- and object-senses in a Fregean thought would be nothing but a pattern of functional applications and abstractions. For example, instead of

for every two integers, there is a third that is greater than their sum, i.e.,  
 $(\forall x)(\forall y)((Ix \ \& \ Iy) \supset (\exists z)(Iz \ \& \ z > x+y)),$

which we might rewrite in functional notation as

$\forall(\lambda x \forall(\lambda y \supset (\&(I(x), I(y)), \exists(\lambda z > (z, +(x, y))))))$ ,

we would have the “form”

$\Phi(\lambda x \Phi(\lambda y \zeta(\theta(\pi(x), \pi(y)), \Psi(\lambda z \kappa(z, \mu(x, y))))))$ .

But at this level of abstraction the *logical* structure of the claim is no longer fully in view, since we’ve “abstracted from” the logical constants. We could continue to maintain that logic is unrestrictedly formal, but only at the cost of confining logic to the meager resources of function application and abstraction, and removing conjunction, negation, quantification, and other notions generally regarded as logical from its purview. It seems more reasonable to follow Frege in denying that logic is unrestrictedly formal.<sup>25</sup>

For Frege, then, logical notions cannot be distinguished from non-logical ones on by *generic* features of their semantic role. Both the logical notion *identity* and the non-logical notion *is taller than* are first level functions with two arguments. They differ only in their graphs:

---

<sup>25</sup>One puzzle: in the long passage quoted from FG2, Frege lists *subsumption* (the falling of an object under a concept or a first level concept under a second level concept, PW:193, 213) as one of the notions proper to logic. But subsumption is not a concept or relation, but the notion of functional application itself (or rather a restriction of that notion, since not all functions are concepts). What is more, subsumption does not correspond to any *symbol* in the Begriffsschrift, so logical relations that depend on it *would* be preserved under arbitrary category-preserving translations. It is therefore hard to see how it belongs with identity, negation, and concept subordination in Frege’s list.

$X$	$Y$	$X$ is taller than $Y$	$X$ is identical with $Y$
George Bush	Jesse Jackson	False	False
Odysseus	Penelope	True	False
George Bush	George Bush	False	True
etc.	etc.	etc.	etc.

Similarly, both the universal quantifier and the second level concept *some philosophers* are second level functions with one argument. The line between logical and non-logical notions must therefore appeal to *specific* features of items within each semantic category (perhaps permutation invariance or something similar). For Kant, by contrast, there is a *generic* difference in semantic functioning between logical and non-logical vocabulary.

Notice that Frege's argument against the 3-formality of logic in FG2 does not depend on his commitment to logical objects and his logicist program (though surely it is *motivated*, at least in part, by these commitments). One could consistently hold that logic has its own contentful concepts (negation, identity, concept subordination, etc.) while denying that *objects* can be given through logical means alone. This observation will be important in section 5.4, when we ask where *we* should stand in the debate between Frege and Kant on the nature of logic, given that most philosophers now reject the idea that logic has its own objects.

### 5.2.6 “The most general laws of truth”

To hold that logic is 1-formal but not 3-formal, as Frege does, is to reject Kant's Thesis. On Frege's view, the laws of logic are *both* constitutive norms for thought as such *and* informative about very general features of the “nature of things.” Frege devotes considerable attention (from 1893 on) to explaining how this can be so.

Frege's main concern is to defuse an argument that would establish Kant's Thesis on quite general grounds (grounds independent of any specifically Kantian doctrines). The argument begins with the observation that “law” is used in two senses, which we might call “normative” and “descriptive:” “In one sense a law asserts what is; in the other it

prescribes what ought to be” (GGZ:xv). A *normative law* prescribes what one ought to do or provides a standard for the evaluation of one’s conduct as good or bad. Public statutes, moral codes, and aesthetic canons are all normative laws. A *descriptive law*, on the other hand, describes counterfactually robust regularities in the order of things. Laws of physics, biology, and geometry are descriptive laws. Whether a law is normative or descriptive is an intrinsic feature of its content: no law can be both. The argument Frege is concerned to rebut simply applies this exclusive categorization to the laws of logic. If these laws are normative for thought as such (1-formal), then they cannot be descriptive: they cannot say anything about how things are. Hence, if they are 1-formal, they must also be 3-formal.

Against this line of thought, Frege wants to have it both ways: he wants to conceive logical laws both as prescribing how one ought to think *and* as saying how things are. He does not think that logical laws are explicitly prescriptive as to their *content* (Ricketts 1996:127). They have the form “such and such is the case,” not “one should think in such and such a way”:

The word ‘law’ is used in two senses. When we speak of moral or civil laws we mean prescriptions, which ought to be obeyed but with which actual occurrences are not always in conformity. Laws of nature are general features of what happens in nature, and occurrences in nature are always in accordance with them. *It is rather in this sense that I speak of laws of truth* [i.e., laws of logic]. Here of course it is not a matter of what happens but of what is. (Th:58)

But although logical laws are not prescriptive in their content, they *imply* prescriptions and are thus prescriptive in a broader sense: “From the laws of truth there *follow* prescriptions about asserting, thinking, judging, inferring” (Th:58, emphasis added). Hence

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

Frege’s line of thought here is subtle enough to deserve a little unpacking. Consider the statement “the white King is at C3.” Though the statement is descriptive in its content, it

has prescriptive consequences in the context of a game of chess: for instance, it implies that white is prohibited from moving a bishop from C4 to D5 if there is a black rook at C5. Now instead of chess, consider the “game” of thinking about the physical world. As in chess, “moves” in this game—thoughts—can be assessed as correct or incorrect. Thoughts about the physical world are correct to the extent that they accurately depict the way the world is. Laws of physics (for instance, Maxwell’s equations) are descriptive laws; they tell us how the physical world *is*. But in the context of the “game” or activity of thinking about the physical world, they have prescriptive consequences: one *ought not* make judgments that conflict with them. In so far as one’s activity is to count as thinking about the physical world, it must be assessable as correct or incorrect (true or false) by reference to the laws of physics.<sup>26</sup> In this sense, the laws of physics provide *constitutive norms* for the activity of thinking about the physical world.

This is not to say that one cannot think *wrongly* about the physical world: one’s thoughts need not *conform to* the norms provided by the laws of physics; they need only be *assessable* in light of them. Nor is it to say that one must be *aware* of these laws in order to think about the physical world. The point is simply that to count someone as thinking about the physical world is *ipso facto* to hold her thoughts assessable for truth or falsity by their agreement or disagreement with the laws of physics. Someone whose thoughts were not so assessable could still be counted as *thinking*, but not as thinking about the physical world. It is in this sense that Frege holds that a law of physics “. . . can be conceived as prescribing that one ought to think in conformity with it, and is thus in that sense a law of thought.”

On Frege’s view, then, laws of physics cannot be distinguished from laws of logic on the grounds that the former are descriptive and the latter prescriptive. Both kinds of laws are descriptive and have prescriptive consequences. They differ only in the activities for which they provide constitutive norms. While physical laws provide norms for thought *about the physical world*, logical laws provide norms for thought *as such*. To count an activity as

<sup>26</sup>There are of course other dimensions of correctness of thought besides truth and falsity: e.g., standards for proper justification.

thought about the physical world is to hold it assessable in light of the laws of physics; to count an activity as thought at all is to hold it assessable in light of the laws of logic. (As in the case of the laws of physics, there is no implication that thought must *conform* to these laws in order to count as thought: the point is that they must be *assessable* in light of these laws.)<sup>27</sup>

Frege often casts the difference between logical laws and laws of the special sciences as a difference in *generality*: logical laws are more general in the sense that they “. . . prescribe universally the way in which one ought to think *if one is to think at all*” (GGZ:xv), as opposed to the way in which one ought to think in some particular domain (cf. PW:1456). Thus, in his 1897 “Logic” manuscript, Frege writes:

Like ethics, logic can also be called a normative science. How must I think in order to reach the goal, truth? We expect logic to give us the answer to this question, but we do not demand of it that it should go into what is peculiar to each branch of knowledge and its subject-matter. On the contrary, the task we assign logic is only that of saying what holds with the utmost generality for all thinking, whatever its subject-matter. (PW:128)

Another way Frege sometimes puts the point is by saying that logic is concerned with *truth*, in the way that chemistry is concerned with *acid* and *alkaline* (among other properties) (PW:128):

All sciences have truth as their goal; but logic is also concerned with it in a quite different way: logic has much the same relation to truth as physics has to weight or heat. To discover truths is the task of all sciences; it falls to logic to discern the laws of truth. (Th:58; cf. GGZ:xvi, PW:3)

One should not be misled by such passages into thinking that logic for Frege concerns itself not with the world itself, but with the relation of thoughts to the world. Frege is not claiming that sentences expressing logical laws will employ an expression for truth, in the way that sentences expressing laws of physics employ expressions for weight and heat. What distinguishes truth from other predicates, Frege claims, “. . . is that predicating it is always

---

<sup>27</sup>See section 3.1, above. Frege notes that the traditional expression “laws of thought” is misleading in that it suggests laws to which all thought conforms, and should thus be avoided in logic, even though there is a legitimate sense in which logical laws *are* laws of thought: PW:145, cf. PW:4.)

included in predicating anything whatever” (PW:129). In making claims about heat, for instance, physics is purporting to tell us *truths* about heat. Thus we can think of the laws of physics as laws of truth about heat, weight, and other physical properties. In general, the laws of a special science X are laws of truth about the subject matter of X. Because logic has no special subject matter, its laws are laws of truth about any subject matter: laws of truth *simpliciter*.<sup>28</sup> Thus, in saying that logic concerns itself with truth in the way that chemistry concerns itself with acid and alkali, Frege is not suggesting that logic has as its subject matter a specifically semantic property with which the special sciences are not concerned. *All* sciences are concerned with truth. To say that truth is the subject matter of logic is just to say that logic has no *special* subject matter.<sup>29</sup>

In this way Frege manages to hang onto the Kantian point that logic is a normative discipline—that it tells us how we *ought* to think, not how we *do* think—while rejecting Kant’s view that logic says nothing about how things *are*. On Frege’s view, the laws of logic are (from different points of view) both the most general laws of reality and the constitutive norms for thought as such.

### 5.2.7 How Frege resists the Kantian argument

We have seen Frege’s grounds for rejecting the Kantian claim that logic is 3-formal. But we have not yet seen in detail how Frege can escape the Kantian argument for that claim (detailed in section 4.4, above). Let us work backwards through the Kantian argument to see exactly where Frege gets off the bus, and on what grounds.

The Kantian argument derives its conclusion,

**(LC)** General logic abstracts entirely from the content of concepts.

from two other claims:

**(LS)** General logic abstracts entirely from the relation of thought to sensibility.

<sup>28</sup>This is essentially the interpretation articulated in Ricketts 1996:127.

<sup>29</sup>Cf. PW:252: “... ‘true’ only makes an abortive attempt to indicate the essence of logic, since what logic is really concerned with is not contained in the word ‘true’ at all, but in the assertoric force with which a sentence is uttered.”

and

(CS) Concepts can have content only in relation to sensibility.

Frege accepts (LS); his disagreement with Kant concerns (CS). What entitles Frege to reject (CS)? (CS) is supported by

(CO) A concept has content (objective validity, significance) only in so far as it applies to some object that could be given to us in intuition (that is, in a singular representation).

and

(OS) Objects can be given to us only in sensibility. That is, for us (as opposed to God), all singular representations are sensible.

Frege rejects both of these principles. Interestingly, he argues against both of them in the FA, *before* making any controversial assumptions about the scope of logic.

### Frege's rejection of (CO)

Recall that Kant holds (CO) because he holds

(CJ) Concepts can be used only in judgment.

and

(JO) Judgment essentially involves the subsumption of an object or objects given in intuition under a concept.

Frege accepts (CJ): that is, he agrees that the significance of concepts is exhausted by their semantic role in propositions (FA:x, cf. §60). But he rejects (JO). His rejection of (JO) is bound up with his claim that ascriptions of number (e.g., “There are two moons of Mars”) assert properties of *concepts*, not objects. Although such ascriptions are objective judgments, he argues, they do not involve the subsumption of any object under a concept.

... the content of a statement of number is an assertion about a concept. This is perhaps clearest with the number 0. If I say “Venus has 0 moons”, *there simply*

*does not exist any moon or agglomeration of moons for anything to be asserted of; but what happens is that a property is assigned to the concept “moon of Venus”, namely that of including nothing under it. (FA:§46, emphasis added)*<sup>30</sup>

If the use of concepts in judgment were limited to the subsumption of objects under them, Frege argues, then judgments of nonexistence (and hence also judgments of existence) would be inexplicable (§49). But such judgments are important in science and mathematics. Thus, a concept *can* have objective content—that is, a use in objective judgments—independently of its “relation to objects,” i.e., of whether an object can be given that falls under it. Frege emphasizes that even self-contradictory concepts, like *rectangular triangle*, have objective content, despite the fact that no object can fall under them, because they can be used in propositions asserting that they have no instances (§53, §74, §94, CES:454, LI:159, RH:326-7, PW:124).<sup>31</sup> Since Frege maintains that concepts have a use in judgment beyond the subsumption of objects under them, he can reject Kant’s argument for (CO), independently of any assumptions about the scope of logic.

### Frege’s rejection of (OS)

Let us now turn to (OS). Frege’s rejection of (OS) is bound up with his construal of numerals as proper names, and hence of numbers as objects:

I must also protest against the generality of Kant’s dictum: without sensibility no object would be given to us. Nought and one are objects which cannot be given to us in sensation. (§89)

This reasoning is pretty compelling, even in advance of the logicist reduction, provided that one agrees with Frege that the numbers are objects. For the numbers (especially 0 and 1) are certainly not objects we can perceive with our senses (FA:§§61, 62). Kant gets

<sup>30</sup>Frege argues that even “all whales are mammals” is not an assertion about any object or objects, but an assertion about the *concepts whale* and *mammal* (§47). Cf. Frege CES:454: “If I utter a sentence with the grammatical subject ‘all men’, I do *not* wish to say something about some Central African chief wholly unknown to me.”

<sup>31</sup>In fact, Frege defines the number 0 (and indirectly the other numbers as well) in terms of the concept *not identical to itself*, under which no object can fall (1884:§74). Compare Kant: “The object of a concept which contradicts itself is nothing, *because the concept is nothing*, is the impossible, e.g. a two-sided rectilinear figure. . .” (KrV:A291/B348, emphasis added).

around this problem not by taking numbers to be objects of the senses, but by denying them objecthood altogether. For Kant, numerals are not names of objects. Arithmetic applies directly to magnitudes given from outside arithmetic, e.g., spatial magnitudes.<sup>32</sup> At the risk of some anachronism, we can think of Kant as construing statements containing numerals (e.g., “ $1 + 2 = 3$ ”) as disguised quantificational statements (“if there is exactly one F and exactly two Gs, and the Fs and Gs are distinct, then there are exactly 3 things that are F or G”).<sup>33</sup> Frege, on the other hand, holds that because numerical terms behave inferentially like names of objects, they are names of objects (§57). Numerical terms can be formed using definite descriptions and used in genuine identity statements that license intersubstitution (1884:§57); they have no plurals (§68 n.); they do not function logically like adjectives (§29-30).

Evidently this dispute hinges on very general philosophical issues about objecthood. On Frege’s view, as Ricketts 1986 has argued, “[o]ur grasp of the notion of an object... is exhausted by the apprehension of inference patterns and the recognition of the truth of the basic logical laws in which [first-level] variables figure” (89). Thus, if numerical terms behave like proper names, then numbers are objects: there is nothing more to be said. For Kant, on the other hand, “object” is defined at the transcendental level, as the locus of objective relations of representations: “that in the concept of which the manifold of a given intuition is *united*” (KrV:B137). Unlike Frege, Kant has a perspective from which he might deny that terms that behave logically as singular terms really denote objects. My point here is not to adjudicate this dispute, however, but to point out that it is in no way a dispute about the scope or nature of *logic*. Frege’s reasons for rejecting (OS) presuppose only his general views about objecthood; they are independent of any premises about logicity that might be disputed by a Kantian. To be sure, Frege appeals to the logicity of his Begriffsschrift in showing *how* numbers can be given to us if not through the senses, but his conviction

<sup>32</sup>Parsons 1969:147-9, Friedman 1992:112-3.

<sup>33</sup>Such sentences can be rendered in first-order logic with identity without the use of numerical singular terms.

that numbers are nonsensible objects does not depend on the resolution of this puzzle.<sup>34</sup>

In sum: the dispute between Kant and Frege about the formality of logic is not merely a verbal disagreement about the *meaning* of “logic,” but a substantive dispute that turns on much more general philosophical questions, namely:

- Can an objective judgment be about a *concept*, or must all judgments involve the subsumption of objects under concepts?
- What is an object? Do numerals denote objects?

Frege’s answers to these questions entitle him to reject the Kantian inference from the generality to the formality of logic, thereby opening up the conceptual space for a general logic through which objects can be given to us and knowledge extended. As I have argued, the grounds Frege offers for his answers do not presuppose the possibility of such a logic. In arguing that arithmetic is reducible to logic, then, Frege is not talking past Kant or “changing the subject.” Despite the large differences in what each takes to be logic and what each believes to follow from logicity, they agree in taking the defining mark of logicity to be its “generality” or 1-formality.

### 5.3 Is Frege’s logic 2-formal?

Our main concern so far has been with Frege’s rejection of Kant’s Thesis: his claim that logic is 1-formal but not 3-formal. But what about 2-formality? Frege’s talk of logic as disregarding the “particular characteristics of things” (BGS:Preface) and as “[transcending] all particulars” (ibid.) certainly suggests something like permutation invariance (see section 3.2, above, and chapter 6, below).<sup>35</sup> And at least one reader (Kit Fine) has suggested explicitly that permutation invariance “. . . is the formal counterpart to Frege’s idea of the

<sup>34</sup>Witness the fact that Frege’s conviction that numbers are nonsensible objects persists long after the collapse of his logicist program (e.g., PW:265, though he expresses some doubt at 263).

<sup>35</sup>Compare FA:4: “For a truth to be a posteriori, it must be impossible to construct a proof of it without including an appeal to facts, i.e., to truths which cannot be proved and are not general, since they contain assertions about particular objects.”

generality of logic” (1998:556). Against this, I will argue that Frege cannot have held that logic is 2-formal, and that we must understand his talk of the generality of logic in terms of its 1-formality.

The argument is simple: if arithmetic can be completely reduced to logic, and the numerals are object-denoting terms, as Frege holds, then logic cannot completely disregard the particular characteristics of objects. Logic must attend to the differences between different numbers: it must be capable of establishing (for instance) that 7, but not 6, is prime.<sup>36</sup> For Frege, the concept *is a prime number* is definable in purely logical terms, even though it is sensitive to the particular identities of objects (that is, it is not permutation invariant).

In the system of the BLA, numbers are defined as the *courses-of-values* of particular functions. The system contains a primitive functor for forming names of courses-of-values from names of functions, as well as a Basic Law (the infamous fifth) allowing the introduction of such names into purely logical proofs. These names are plainly names for particular objects. But that is not all. For Frege, *every sentence* is the name of a particular object: a truth value. And because the truth values are objects, not even the *truth functions* in Frege’s logic are insensitive to differences between particular objects. Negation and the conditional must be able to distinguish the True from all other objects. Finally, every one of Frege’s logical laws employs a concept, the “horizontal” ( $\text{—}$ ), whose extension is {the True} (BLA:§5). The horizontal is plainly no more permutation-invariant than the concept “is identical with Ronald Reagan,” whose extension is {Ronald Reagan}.

What, then, of Frege’s claims in the Preface of the BGS that logic disregards “particular characteristics of things” and “transcends all particulars”? These words are not incompatible with anything he says in the BGS, which does not posit “logical objects.” And I have argued (section 5.1) that in the BGS, Frege has not yet fully liberated himself from a Kantian conception of logic. It seems reasonable, then, to take these claims as expressions of

<sup>36</sup>In arguing against inductive justifications of arithmetic, Frege emphasizes that each number “has its own unique particularities” (FA:§10).

an immature view of logic that Frege later abandons.

It is true that even in his mature works, Frege characterizes logic by its maximal generality. But this fact need not cause any puzzlement, provided we understand “generality” as 1-formality, rather than 2-formality (see section 3.5, above). Logic is general in the sense that it is *generally applicable* to all thought *as such*, not in the sense that it abstracts completely from the particular identities of objects. If we confuse these senses of generality, it will look puzzling how Frege can hold both

1. Arithmetic is reducible to logic, and
2. Arithmetic refers to and distinguishes between particular objects (the numbers).

For example, Sluga 1980 asks how propositions “that make assertions about particular numbers ... [or] assert the existence of numbers ... could be regarded as universal, and therefore logical, truths” (109). He resolves the tension by arguing that Frege’s Basic Law V is not “irreducibly existential,” so that Frege’s logic is “non-ontological in character” (110). But whether or not this solution can be made out,<sup>37</sup> it is *unnecessary* once we distinguish the two senses of logical “universality.” The singular reference of arithmetical propositions *does* count against their 2-formality, but it does not count against their 1-formality. As Sluga himself points out, “[t]here is another and deeper sense in which [logical] laws can be called universal. They are universal because they are universally applicable” (111).

## 5.4 Does Russell’s Paradox vindicate Kant’s Thesis?

In section 5.2, I argued that Frege’s explicit rejection of Kant’s Thesis was *motivated* by his recognition (starting in 1884) of logical objects. The technical results of the BGS demonstrated that much more can be done with pure general logic than Kant had thought,

---

<sup>37</sup>I find it implausible because, although Basic Law V does not contain an existential quantifier (i.e., a negated universal quantifier), that is no grounds for claiming that it is “universal” as opposed to “irreducibly existential,” as Sluga suggests. In Frege’s logic, the use of a functor like the smooth breathing operator in Basic Law V already involves existential commitment: it is presupposed that all singular terms formed by applying the functor to a concept expression refer.

but they did not by themselves demand a rejection of Kant's Thesis. Before the introduction of logical objects, it was still possible to hold on to the Kantian equation

$$\text{semantic content} = \text{intuitive content}$$

and to conceive of logic as abstracting entirely from the content of concepts and judgments, articulating only their schematic forms. Indeed, as I have argued, Frege seems to have waffled between a Kantian conception of logic and something akin to his own later view. It was the introduction of logical objects that *demand*ed a reassessment of the Kantian conception of logic and the explicit articulation of an alternative. For Kant, logic abstracts from all relation to objects, but Frege's Basic Law V directly implies the existence of objects that can be given by logical means alone (extensions, or more generally courses-of-values).

However, as Russell showed, Basic Law V is inconsistent with the other axioms of Frege's logic. After several failed attempts to patch up his theory of extensions, Frege eventually comes to reject the notion of *extension* entirely. In 1924/5, he writes that "the extension of the concept *a*" is a proper name to which no object corresponds (PW:269). He abandons his logicism and tries to reconstruct arithmetic on a *geometrical* basis (PW:277, 279). And he declares that the "logical source of knowledge" "...on its own cannot yield us objects" (PW:279).

In conceding that logic alone "cannot yield us objects," is Frege giving up his entitlement to reject Kant's Thesis? Does the abandonment of "logical objects" mean a return to a Kantian conception of logic? Peter Hylton suggests as much:

For Kant, logic has no objects of its own, and does not even deal with objects; its concern is with the understanding and its form. . . . Russell's propositions and propositional functions, by contrast, are logical objects (as are Frege's *Wertverläufe*.) One way to understand the significance of Russell's paradox and related paradoxes, is as showing that Kant was right on this issue, and Russell and Frege wrong. (Hylton 1990a:170 n. 39)

But Frege's abandonment of extensions does not force him to revert to the Kantian view that logic abstracts from all semantic content. One reason is that extensions are not the *only* non-sensible objects he acknowledges. As late as 1919 (PW:255), Frege uses the

definite descriptions “the *True* and the *False*,” which suggests that he still takes them to be objects, even though he has come to have doubts about extensions.<sup>38</sup> If Frege had really rejected the view that the truth values are objects—a central tenet of his semantic theory—one would have expected some acknowledgement of this change. Moreover, Frege still takes *thoughts* (and senses in general) to be non-sensible objects (or, if he doesn’t, he doesn’t tell us). Thus, Frege’s claim that logic “. . . on its own cannot yield us objects” does not prove conclusively that he has abandoned logical objects altogether: perhaps what he means is that logic alone cannot yield the sorts of objects we would need in order to do arithmetic.

But even if Frege did completely repudiate objects beyond those given in sensible intuition, he could still avoid taking logic to be 3-formal. As I have argued in section 5.2.5, Frege’s argument that logic is not 3-formal does not depend on the premise that logic can yield singular reference to objects and non-trivial existence claims. The crux of the argument is that logic has its own *concepts* and *relations* (identity, negation, subsumption, etc.), and that these concepts and relations have content in just the same sense as non-logical concepts and relations do: they are functions from objects to truth values (or, in the case of second level concepts, from functions to truth values). Frege’s logical concepts relate to objects just as directly as concepts of the special sciences. For Kant, by contrast, logic deals only with the *forms* by which concepts are unified in judgment: it abstracts from *all* conceptual content, not just from the contents peculiar to the special sciences.

Even from *Kant’s* perspective, logic’s lack of proprietary objects would not suffice to show that it lacks content. Kant regards arithmetic, algebra, and geometry as contentful sciences, even though he does not countenance specifically mathematical objects (such as shapes or numbers). He makes it clear that construction in pure intuition provides only the *forms* of empirical objects, not distinct mathematical objects as the Platonist would conceive them (KrV:B147, A239-40/B298-9, A224/B271, Thompson 1972-3:338-342, Friedman 1992:101). The mathematical sciences are contentful for Kant not because they have

---

<sup>38</sup>It is true that in the GGZ, the truth values are stipulated to be identical to particular extensions (§10) in order to avoid indeterminacy. But these stipulations are arbitrary: Frege could still have taken the truth values to be objects even after rejecting extensions.

their own objects—they do not—but because their proprietary *concepts* (e.g., the concept of magnitude) are contentful. Mathematics relates to objects indirectly, through *a priori* concepts whose relation to objects is secured in the transcendental deduction (KrV:A239-40/B298-9). Thus, what makes logic purely formal is not that there are no specifically logical objects, but that there are no specifically logical *concepts*.

Thus, although Frege's abandonment of logicism and logical objects makes his position *consistent* with the Kantian claim that logic is 3-formal, it in no way *commits* him to accepting it. On the contrary, his central reasons for rejecting the 3-formality of logic remain unscathed. If we are to accept the claim that logic is 3-formal, we need to be convinced not just that there are no logical objects, but that logic has no contentful concepts or relations.

## 5.5 Frege's influence

In chapter 4, I argued that logical hylomorphism in its original form—the doctrine that logic is 3-formal—is a Kantian innovation, one that is intimately tied up with more characteristically Kantian doctrines, such as the doctrine that semantic content requires the cooperation of both thought and sensible intuition. Our examination of Frege has provided further support for this thesis. As we have seen, Frege can resist Kant's Thesis—the doctrine that logic, if it is to be 1-formal, must also be 3-formal—only because he rejects Kantian assumptions about concepts and objects. The clarity with which Frege articulates his very unKantian philosophy of logic reveals just how optional that philosophy of logic is.

Given Frege's influence on later analytic philosophy, one might have expected his rejection of logical hylomorphism to bury it forever. But oddly, Frege's philosophy of logic proved much less enduring than his logic. Many of the workers who used Frege's technical tools in philosophically-motivated projects (e.g., Carnap) came out of the neo-Kantian tradition and did not take well to Frege's characterization of logic as a substantive science. They were also heavily influenced by Wittgenstein, who developed his own version of logical hylomorphism. Frege's alternative just dropped out of the scene.

A little potted history will help show what I mean. As is well known, the positivists sought to combine Kant's view that logic is 3-formal with the logicist reduction of mathematics to logic. In this way, they hoped to salvage the *a prioricity* of mathematics without having to swallow the rationalist pill of *substantive* (or synthetic) *a priori* knowledge. They could be empiricists without sacrificing the rationalist insight that mathematical knowledge is fundamentally different from empirical knowledge. But unlike Kant, the positivists were not content to understand the 3-formality of logic in terms of "forms of judgment" or "modes of combination" characteristic of the human understanding. They wanted to be more tolerant, acknowledging the possibility of other forms of judgment, and they wanted to explain how we could know logic *a priori*. So they became conventionalists: they held that logical truths were true by convention or stipulation, and that stipulating that these sentences were true fixed the meanings of the terms contained in them.

Certainly one of the turning points of twentieth century philosophy is Quine's (and others') demolition of conventionalism. But notice what happens to logic. One might have expected a return to a more Fregean conception of logic, a conception that rejects 3-formality but retains 1-formality and thereby secures a principled distinction between logic and non-logic. But on Quine's conception, logic has no principled essence at all. His demarcation of logic is thoroughly pragmatic (in the sense of section 1.3.1, above). Quine simply *rejects* the rationalist insight that motivated the positivists: the thought that logico-mathematical knowledge and empirical knowledge are fundamentally different. Logic, mathematics, and empirical science are, on Quine's view, just different regions on a continuum.

In this way, the history of twentieth century analytic philosophy has bestowed on us a Manichean vision of the options in the philosophy of logic: on the one hand, an untenable conventionalism that tries to hang on to logical hylomorphism; on the other, a holistic naturalism that sees only a pragmatic distinction between logic and other disciplines. Either logic is demarcated by its 3-formality, or it has no principled essence at all.

In the din and smoke of the great battle between Quine and the conventionalists, Frege's philosophy of logic has gotten lost. But it may provide just what we need: a moderate middle

ground between logical hylomorphism in its conventionalist guise and Quinean pragmatism. Frege's view makes room for a difference in principle between logic and non-logic without the claim (difficult to defend outside a Kantian context) that logic abstracts entirely from semantic content. Logical truths and inference rules are defined by their fundamental role as constitutive norms for thought. To be sure, it is far from clear that this core idea, common to Kant and Frege, can be made out in a contemporary framework. My point here is that it hasn't been tried—and deserves to be.

In chapter 6, I show one way in which 1-formality might be brought to bear in a contemporary project for demarcating logic.

## Chapter 6

# PERMUTATION INVARIANCE AND LOGICALITY

Of the three notions of logical formality we distinguished in chapter 3, 3-formality and 1-formality have been the most important historically in debates about the demarcation of logic. 2-formality has had little historical importance, except insofar as it is connected with the other two notions. As we have seen, Kant did not regard 2-formality as *sufficient* for logicality (section 4.1.2, above), and Frege did not even regard it as *necessary* (section 5.3, above).

In the contemporary philosophy of logic, by contrast, 2-formality is often taken to be central in discussions of logicality or logical “formality.” A number of philosophers have suggested that logical notions are characterized by their insensitivity to the particular identities of objects (Mautner 1946; Mostowski 1957:13; Tarski 1966; Scott 1970:161; Dummett 1973:22 n; McCarthy 1981; van Benthem 1989; Sher 1991 and 1996, Shapiro 1998:99). The logical notions, on this kind of account, are the notions that are invariant over arbitrary permutations (or, more generally, bijections) of the domain of objects, or those that can be

defined using invariant notions.<sup>1</sup> For example, whereas the quantifier<sup>2</sup> *all chickens* is sensitive to the difference between chickens and cows, the cardinality quantifiers are insensitive to the particular natures of the objects to which they apply: “there are at least five Fs” can be true whether the Fs are numbers, people, places, or diamonds, provided there are at least five of them. Hence the cardinality quantifiers are logical or formal, while the various farm-animal quantifiers are not. (See section 3.2, above.)

As we have seen (section 3.5), permutation invariance provides a way of spelling out the “topic-neutrality” or “maximal generality” of logic. Compared with the other ways of glossing topic-neutrality—3-formality and 1-formality—it is clear and mathematically tractable. Moreover, it gives the “right yield” of logical notions: the notions it certifies as logical (at least in extensional quantificational languages) are just those generally recognized as logical by model-theorists. Sher 1996 notes that the class of logical systems generated by her criterion “partly coincides” with the class of systems studied under the head of “model-theoretic logics” or “abstract logics” (679). On the other hand, the invariance criterion excludes the set-theoretic membership relation, numbers and arithmetical relations, and mereological notions. The claim that the invariance criterion certifies the “right” notions as logical is significantly bolstered by a pair of technical results. Lindenbaum and Tarski 1936 show that all of the notions defined in *Principia Mathematica* are permutation-invariant. Moving in the other direction, McGee 1996 shows that every permutation-invariant operation can be defined by a combination of operations with an intuitively logical character (identity, substitution of variables, finite or infinite disjunction, negation, and finite or infinite existential quantification).

Part of the appeal of the invariance criterion to twentieth century philosophers comes from the success of Klein’s use of invariance to delineate different geometries. In his Erlangen

---

<sup>1</sup>I follow Tarski 1966 in talking of the “logical notions”—the semantic values of logical constants—rather than the “logical constants” (which are interpreted linguistic expressions). Not all of the philosophers I have listed concur in this approach, which is motivated in section 6.3, below.

<sup>2</sup>I use “quantifier” and “operator” ambiguously throughout this chapter, to mean (1) the interpreted linguistic expressions “ $\forall$ ”, “L”, and so on, and (2) the semantic values of these expressions. It should always be clear which sense is at issue.

program (Klein 1893), Klein shows that the notions of Euclidean geometry are invariant under similarity transformations, those of affine geometry under affine transformations, and those of topology under bicontinuous transformations. Tarski 1966 suggests that the logical notions are just those that are invariant under the widest possible group of transformations: the group of *permutations* of the elements in the domain (149; cf. Mautner 1946). Seen in this way, the logical notions are the end point of a chain of progressively more abstract notions defined by their invariance under progressively wider groups of transformations of a domain.

As an account of the distinctive generality of logic, then, permutation invariance has much to recommend it. It is motivated and mathematically precise, it yields results that accord with common practice, and it gives determinate rulings in some borderline cases (for example, set-theoretic membership). Best of all, it offers hope for a sharp and principled line between logic and non-logic that steers clear of the dubious notion of 3-formality and is compatible with Quine's criticisms of conventionalism. It is not surprising, then, that the invariance account is being endorsed more and more frequently by philosophers who need a precise necessary condition for logicity.

In this chapter, I will argue that the permutation invariance criterion does not deliver what it promises. I am not going to deny that invariance under all permutations of objects is in fact a necessary condition for logicity. Indeed, I am going to propose a criterion for logicity with permutation invariance at its very core. The problem with the permutation invariance account as it is usually presented is not that it gives the wrong output, but that it depends crucially on assumptions for which it provides no justification—no doubt because they are hidden in the case to which the account is usually applied, classical extensional logic. I aim to make these assumptions explicit by showing how the account can be extended to multivalued and intensional logics, and to provide a framework within which the assumptions can be justified. Invariance will serve, on this account, as a vehicle for turning questions about logicity into more general questions in the philosophy of language—precisely the kind of questions, I think, on which a demarcation of logic ought

to depend. The resulting account will be recognizable as a version of 1-formality.

The plan for the chapter is as follows. In section 6.1, I formulate the permutation invariance criterion in a functional type theory-theoretic semantics. The notion of logicality defined is actually *presemantic*: it sorts logical from non-logical semantic values in each type, without reference to linguistic expressions (“logical constants”) or their relation to these semantic values. In section 6.2, I discuss the relation between logical notions and logical constants, and I propose three strong necessary conditions for an expression to be a logical *constant*. Because of their sensitivity to the *domain* of an interpretation, quantifiers do not at first appear to meet these conditions. In section 6.3, I show how the quantifiers can be shown to meet the conditions, provided we add some complexity to the type-theoretic presemantics. (Those who just want to get the main point of the chapter may skip this section.)

In sections 6.4 and 6.5, I show how the invariance criterion can be extended to multi-valued and intensional logics. There is a natural and well-motivated way to accomplish this extension, but it requires that we recognize some structure on the basic semantic types as “intrinsic” and consider only permutations of the types that preserve their intrinsic structure. In these cases, it is very clear that the invariance criterion only gives a definite verdict on logicality *relative to* a choice of intrinsic structure for the basic semantic types. By itself, it gives no guidance about what structure should count as intrinsic, and hence no verdict on logicality.

This lacuna in the account has escaped notice, I suggest, only because in the special case to which the invariance criterion has most often been applied—two-valued extensional languages—the appropriate intrinsic structures on both the basic types (Objects and Truth-Values) are degenerate cases. The intrinsic structure on Objects is the null structure, so that *every* permutation is permitted, while the intrinsic structure on Truth-Values (the ordering  $\text{False} \leq \text{True}$ ) admits *no* permutations. Thus in this special case, the invariance criterion can be formulated without explicit reference to structure. But such formulations still *presuppose* a certain choice of structure, and that choice must be justified. The invariance criterion no

more gives an absolute verdict about logicality for classical extensional languages than it does for multivalued or intensional languages (section 6.6). By itself, it does not rule out the set-theoretic membership relation or the mereological sum operator as logical notions.

In section 6.7, I make a suggestion about how we might understand intrinsic structure, and I apply my proposal to a number of test cases (multivalued logics, modal logics, tense logic, extensional logic).

## 6.1 Permutation invariance

My goal in this section is to give a precise definition of permutation invariance for notions of every semantic type used in the semantics for two-valued extensional languages (objects, functions, concepts, relations, truth functions, first- and second-order quantifiers, etc.). In order to achieve the desired generality, I work in a functional type-theoretic framework. Since types are defined recursively in this framework, a single recursive definition of permutation invariance suffices for all types.

I should emphasize that the definition offered in this section applies in the first instance to logical *notions*—that is, items within each semantic type—and not to logical *constants* (that is, *expressions* or grammatical *modes of combination*). Indeed, it says nothing about linguistic expressions or their relation to the notions within each type and is thus properly speaking *presemantic*.<sup>3</sup> The advantages and limitations of this approach will be addressed in the next section.

Before presenting the permutation invariance criterion, it will be necessary to explain the categorial grammar and typed presemantics I am using, and to show how these can be exploited to obtain an elegant compositional semantics for a language. Readers familiar with these ideas may skip directly to section 6.1.4.

---

<sup>3</sup>The term (and the concept) are due to Nuel Belnap (UCL).

### 6.1.1 Categorical grammar

The typed presemantics I will employ is a natural choice for constructing a semantic theory for a language with a categorical grammar. In a categorical grammar,<sup>4</sup> the grammatical categories are defined functionally in terms of a few *basic categories*: for example, T (singular term) and S (sentence). For any two categories  $x, y$ , there is a *derived category*  $(x, y)$ . Expressions in category  $(x, y)$  combine with expressions in category  $x$  to form expressions in category  $y$ . Thus, a one-place predicate has category  $(T, S)$ , since it yields a sentence given any singular term; similarly, a unary quantifier has category  $((T, S), S)$ , since it yields a sentence given any predicate; and a unary sentential operator has category  $(S, S)$ , since it yields a sentence given any sentence.  $N$ -place predicates (and similarly  $n$ -ary quantifiers,  $n$ -ary sentential operators, and so on) can be dealt with in either of two ways: we can either introduced a mechanism for forming a new kind of derived category consisting of *ordered  $n$ -tuples* of elements from existing categories, and take an  $n$ -place predicate as an expression that yields an S when applied to an ordered  $n$ -tuple of Ts; or we can take  $n$ -place predicates as expressions that yield  $(n-1)$ -place predicates when applied to a T. Since the latter course yields a simpler grammar and works just as well for most purposes, I will adopt it in what follows.<sup>5</sup>

Given a specification of the basic categories, all that is required to complete a categorical grammar is an assignment of the language's lexical primitives to categories and a recursive definition of well-formedness. In standard first-order languages, this amounts to assigning “&” to  $(S, (S, S))$ , “F” to  $(T, S)$ , “a” to T, “ $\forall$ ” to  $((T, S), S)$ , and so on, then defining an expression as *well-formed* if and only if

(G1) it is a lexical primitive, or

(G2) it is the concatenation of two well-formed expressions  $x, y$  (in that order)

---

<sup>4</sup>The term is due to Ajdukiewicz. I learned it from Lewis 1970.

<sup>5</sup>The other approach is more appropriate if one wishes to include branching quantifiers and other “independence-friendly” notions in the language (see Hintikka 1998:15). Since this issue is skew to my main concerns in this chapter, I ignore it for now.

such that the category of  $x$  is  $(w, v)$  and the category of  $y$  is  $w$  for any categories  $w, v$  (basic or derived). (The category of the new expression is  $v$ .)

Grammars of this sort produce a parentheses-free “Polish” notation. Thus, for example, we write “ $\sim\&pq$ ” instead of “ $\sim(p\&q)$ .”<sup>6</sup>

In order to handle quantification in a categorial grammar, we will need a  $\lambda$ -abstraction operator.<sup>7</sup> To see why, consider the sentences

- (a) Everyone sleeps.
- (b) Someone sleeps.
- (c) Joe loves Cindy.
- (d) Everyone loves someone.

Let it be a requirement of our analysis that the quantifiers “everyone” and “someone” have the same grammatical category in (a) and (b) as they do in (d), and that “loves” has the same grammatical category in (c) as it does in (d). In (a) and (b) the quantifiers are clearly in the category  $((T,S), S)$ , since they take a one-place predicate and make a sentence; and in (c) “loves” is clearly in the category  $(T, (T,S))$ .<sup>8</sup> But how are we to understand (d)? If “everyone” is a  $((T,S), S)$ , then “loves someone” must be a  $(T,S)$ —that is, a one place predicate. And so it is: by adding a singular term to it, we obtain a sentence (e.g., “Joe loves someone”). Thus, if “loves” is a  $(T, (T,S))$ , “someone” must be a  $T$ —or alternatively a  $((T,(T,S)),(T,S))$ . It cannot be a  $((T,S), S)$ , as it must be if “someone” is not to be ambiguous between (b) and (d).

What is needed in order to avoid this problem is a way of making a  $(T, S)$  “loves someone” out of a  $(T, (T,S))$  “loves” and a  $((T,S), S)$  “someone.” The  $\lambda$ -abstraction

<sup>6</sup>One reason for using this kind of grammar—in which the only “mode of combination” is concatenation—is that it makes us less likely to be “grammatical chauvinists” than grammars in which there is a mode of combination for each logical constant (see section 2.3.3, above). The strategy is to put logical and non-logical expressions on a par, grammatically, so that we can better see how they differ *semantically*.

<sup>7</sup>In effect, we are moving from what Cresswell 1973 calls “pure categorial languages” to what he calls “ $\lambda$ -categorial languages.” My whole discussion here (including the choice of examples) is indebted to Cresswell (esp. 1973:81-3).

<sup>8</sup>An alternative approach (Montague’s) is to put names in same category as quantifiers— $((T,S),S)$ —and take “loves” as a  $((T,S),S), ((T,S),S), S)$  (see Cresswell 1973:81).

operator provides such a way. First, we use variables in category  $T$  to form the  $S$  “loves  $x$   $y$ ” (in English, “ $x$  loves  $y$ ”). Then we form the abstract “ $\lambda y$  loves  $x$   $y$ ” (in English, “is loved by  $x$ ”), which is in category  $(T, S)$ . To this we can apply the quantifier “someone” to get the  $S$  “someone  $\lambda y$  loves  $x$   $y$ ” (in English, “someone is loved by  $x$ ”). Finally, we form the abstract “ $\lambda x$  someone  $\lambda y$   $x$  loves  $y$ ” (in English, “loves someone”), which is the desired  $(T, S)$ . Sentence (d) comes out as “Everyone  $\lambda x$  someone  $\lambda y$  loves  $x$   $y$ .”

The addition of a  $\lambda$ -abstraction operator to the grammar requires some minor changes. In the lexicon, we distinguish between *variables* and *simple terms*. And in the definition of well-formedness, we add a third clause: an expression is well-formed if

- (G3)** it is the concatenation of  $\lambda$ ,  $x$ , and  $y$  (in that order) such that  $x$  is a variable and  $y$  is a well-formed expression. (The category of the new expression is  $(w, v)$ , where  $w$  is the category of  $x$  and  $v$  is the category of  $y$ .)

The grammar now has two modes of combination: concatenation and  $\lambda$ -addition. The abstraction operator  $\lambda$  will be used to handle variable binding in the semantics.<sup>9</sup>

### 6.1.2 Typed presemantics

The natural semantics for a categorial grammar draws semantic values for expressions from a typed presemantics. Corresponding to the basic and derived categories in a categorial grammar are basic and derived *types* in a typed presemantics (or “type theory,” as it is usually called). For example, corresponding to the basic categories  $T$  and  $S$ , we might have the basic types  $O$  and  $V$  (for “object” and “truth value”). We can think of these basic types as *sets* (of objects and truth values, respectively). Corresponding to each derived category  $(w, v)$ , there is a derived type, the set of *functions* from the type corresponding

<sup>9</sup>It would, of course, be possible to treat abstraction operators as generic items in the lexicon (as Lewis 1970 does), rather than singling them out for special treatment in the grammar. Abstraction operators on individual variables would be items in the category  $(S, (T, S))$ , abstraction operators on predicate variables would be items in  $(S, ((T, S), S))$ , and so on. But because abstraction will play a special role in the semantics (as the only mechanism for variable binding), it is useful to treat it differently in the grammar as well (after the fashion of Cresswell 1973). In the reprint version of Lewis 1970, Lewis asserts that he now thinks Cresswell’s approach is superior (Lewis 1983:231).

to category  $w$  to the type corresponding to category  $v$ . Thus, the type corresponding to the category  $(T, S)$  is the set of functions from  $O$  to  $V$  (i.e., Fregean “concepts”); the type corresponding to  $(S, S)$  is the set of functions from  $V$  to  $V$  (i.e., truth functions); the type corresponding to  $((T, S), S)$  is the set of functions from Fregean concepts to truth values (i.e., unary quantifiers); and so on. We name these derived types by giving their function spaces: for example, the type corresponding to the category  $(T, S)$ —i.e., the set of functions from  $O$  to  $V$ —is called “ $(O \Rightarrow V)$ .”

### 6.1.3 Compositional semantics

If we let the semantic values for expressions in a category be the items in the corresponding type, we obtain an elegant compositional semantics: an account of how the semantic values of complex well-formed expressions of the language are determined by the semantic values of their parts.

First, let the semantic values of lexical primitives be given by an *interpretation function*  $i$  (for constant terms) and a *variable assignment*  $z$  (for variables). The function  $i$  takes constant terms to items in the types corresponding to the terms’ grammatical categories (e.g., it takes terms in  $(T, S)$  to items in the type  $(O \Rightarrow V)$ .) Similarly, the function  $z$  takes variable terms to items in the types corresponding to the variables’ grammatical categories (e.g., it takes first-level variables to items in  $O$ ).

Given  $i$  and  $z$ , we can compute the semantic value of an arbitrarily complex expression  $x$ ,  $\|x\|_i^z$ , in the familiar recursive way.

- If  $x$  is a primitive constant term, then  $\|x\|_i^z = i(x)$ .
- If  $x$  is a primitive variable term, then  $\|x\|_i^z = z(x)$ .
- If  $x = 'mn'$ ,<sup>10</sup> where  $m$  and  $n$  are well-formed expressions in the categories  $(u, v)$  and  $u$ , respectively, for any categories  $u$  and  $v$ , then  $\|x\|_i^z =$  the value of  $\|m\|_i^z$  on  $\|n\|_i^z$ .

<sup>10</sup>I use ‘ ’ for corner-quotes. Thus, ‘ $mn$ ’ is the concatenation of  $m$  and  $n$ .

- If  $x = \lambda mn$ , where  $m$  is a variable in category  $u$  and  $n$  is a well-formed expression in category  $v$ , for any categories  $u$  and  $v$ , then  $\|x\|_i^z =$  the function  $f$  from the type corresponding to  $u$  to the type corresponding to  $v$  such that for all  $w$ ,  $f(w) = \|n\|_i^{[w/m]z}$ , where the assignment  $[w/m]z$  is the same as  $z$  except possibly on the variable  $m$ , and  $[w/m]z(m) = w$ .

Since there are only two modes of combination in the grammar, these clauses suffice to determine the semantic values of every well-formed expression in the language.

By construing quantifiers grammatically as applying directly to (abstracted) (T,S)'s, instead of binding variables themselves, we relegate all of the variable binding in the semantics to the  $\lambda$ -abstraction operator.<sup>11</sup> In particular, we do not need separate recursive clauses for the quantifiers (as we do in more orthodox Tarski-style semantics). This way of proceeding makes sense, given our theoretical goals, because it allows us to see the semantic roles of quantifiers as exhausted by their semantic values.<sup>12</sup> We can then consider the permutation invariance criterion for logicality entirely at the presemantic level: that is, at the level of the semantic values themselves, independently of any consideration of their relations to linguistic entities.

#### 6.1.4 The permutation invariance criterion

We are now in a position to give a precise formulation of Tarski's (1966) suggestion that "... we call a notion 'logical' if it is invariant under all possible one-one transformations of the world onto itself" (149). Suppose the basic grammatical categories of the language are T (singular term) and S (sentence), and let the corresponding basic types be the sets O (of objects) and V (of truth values). "One-one transformations of the world onto itself" can then be construed as permutations of the set O, and "notions" as items in basic or derived presemantic types. In order to define invariance for arbitrary types, we need to define the way in which a permutation of O affects an arbitrary type derived from O.

<sup>11</sup>The technique is due to Church 1940.

<sup>12</sup>Ah, but musn't the quantifiers still receive different semantic values in models with different domains? This objection will be considered in section 6.2.3 and answered in section 6.3.

**Induced Transformation.** Where  $\sigma$  is a permutation of  $O$ ,  $\sigma^Z$  is the *transformation on the type  $Z$  induced by  $\sigma$* . It is defined inductively on the complexity of  $Z$ :

- if  $Z=O$ , then for all  $w \in Z$ ,  $\sigma^Z(w) = \sigma(w)$
- if  $Z=V$ , then for all  $w \in Z$ ,  $\sigma^Z(w) = w$
- if  $Z=(X \Rightarrow Y)$ , for any types  $X, Y$ , then for all  $w \in Z$ ,  $\sigma^Z(w) = \sigma^Y \circ w \circ (\sigma^X)^{-1}$ .

The inductive clause is perhaps easier to understand with a picture:

$$\begin{array}{ccc}
 X & \Longrightarrow \sigma^X \Longrightarrow & X \\
 \Downarrow w & & \Downarrow \sigma^Z(w) = \sigma^Y \circ w \circ (\sigma^X)^{-1} \\
 Y & \Longrightarrow \sigma^Y \Longrightarrow & Y
 \end{array}$$

Permutation invariance can then be defined as follows:

**Permutation Invariance.** A semantic value  $w$  in type  $Z$  is *permutation-invariant* iff for every permutation  $\sigma$  of  $O$ ,  $\sigma^Z(w) = w$ .

And Tarski’s suggestion amounts to this:

**Invariance Criterion.** A semantic value  $w$  in type  $Z$  is *logical* iff it is permutation-invariant.

Let’s look at which items in each type count as logical on Tarski’s criterion. Trivially, everything in  $V$ —as well as the truth-functions in  $(V \Rightarrow V)$ ,  $(V \Rightarrow (V \Rightarrow V))$ , and so on—counts as logical. Equally trivially, no item in  $O$  is logical, since for any  $w \in O$ , there is a permutation  $\sigma$  such that  $\sigma(w) \neq w$ . It is only when we get to the more complex categories that the criterion becomes non-trivial. For example, let  $z \in (O \Rightarrow V)$  be the function that takes every object in  $O$  to True. Then for any permutation  $\sigma$ ,  $\sigma^{(O \Rightarrow V)}(z) (= z \circ \sigma^{-1})$  still

takes every object in  $O$  to True; that is,  $\sigma^{(O \Rightarrow V)}(z) = z$ . Only one other item in  $(O \Rightarrow V)$  is permutation-invariant: the function that takes every object in  $O$  to False.

When we come to  $((O \Rightarrow V) \Rightarrow V)$ , there are many invariant items. Let  $y$  be the function that takes every item in  $(O \Rightarrow V)$  which is true of exactly five objects to True, and everything else in  $(O \Rightarrow V)$  to False. (This is the semantic value of the quantifier “there are exactly five things such that...”.) Then  $\sigma^{((O \Rightarrow V) \Rightarrow V)}(y) = y$ . To show this, we’ll need an easy theorem:

**Inverse theorem.** For any type  $Z$ ,  $(\sigma^{-1})^Z = (\sigma^Z)^{-1}$ .<sup>13</sup>

Using this theorem, we can show that for any  $w \in (O \Rightarrow V)$ ,  $\sigma^{((O \Rightarrow V) \Rightarrow V)}(y)(w) = y(w \circ \sigma)$ :

$$\begin{aligned} \sigma^{((O \Rightarrow V) \Rightarrow V)}(y)(w) &= \\ [\sigma^V \circ y \circ (\sigma^{(O \Rightarrow V)})^{-1}](w) &= \\ [y \circ (\sigma^{(O \Rightarrow V)})^{-1}](w) &= \\ [y \circ ((\sigma^{-1})^{(O \Rightarrow V)})](w) &= \\ y((\sigma^{-1})^{(O \Rightarrow V)}(w)) &= \\ y((\sigma^{-1})^V \circ w \circ ((\sigma^{-1})^O)^{-1}) &= \\ y((\sigma^V)^{-1} \circ w \circ ((\sigma^O)^{-1})^{-1}) &= \\ y(w \circ \sigma) & \end{aligned}$$

Since  $\sigma$  is a permutation of  $O$ ,  $(w \circ \sigma)$  will take the value True on exactly as many objects as  $w$  will. So  $y(w \circ \sigma)$  will be true iff  $y(w)$  is true. Thus,  $\sigma^{((O \Rightarrow V) \Rightarrow V)}(y)(w)$  will be true iff  $y(w)$  is true. That is,  $\sigma^{((O \Rightarrow V) \Rightarrow V)}(y) = y$ .

As a final example, consider the unary identity function ( $i \in (O \Rightarrow O)$ ) that takes each object in  $O$  to itself. Since  $\sigma^{(O \Rightarrow O)}(i) = \sigma \circ i \circ \sigma^{-1} = \sigma \circ \sigma^{-1} = i$ ,  $i$  is also a logical semantic value according to Tarski’s criterion.

As an example of a semantic value that *fails* to meet the criterion, consider the function  $f \in (O \Rightarrow O)$  that takes each object to the singleton set containing it. Let  $\sigma$  be the

<sup>13</sup>Proof is by induction on the complexity of  $Z$ . The inductive step goes like this: let  $\sigma^{(X \Rightarrow Y)^{-1}}(f) = g$ . Then  $\sigma^{(X \Rightarrow Y)}(g) = f$ , so  $\sigma^Y \circ g \circ (\sigma^X)^{-1} = f$ , so  $g \circ (\sigma^X)^{-1} = (\sigma^Y)^{-1} \circ f$ , so  $g = (\sigma^Y)^{-1} \circ f \circ \sigma^X = (\sigma^Y)^{-1} \circ f \circ ((\sigma^X)^{-1})^{-1} = (\sigma^{-1})^Y \circ f \circ ((\sigma^{-1})^X)^{-1} = (\sigma^{-1})^{(X \Rightarrow Y)} f$ .

permutation that takes each integer to its successor and leaves all other objects the same. Then  $f(1) = \{1\}$ , but  $\sigma^{(O \Rightarrow O)}(f)(1) = \sigma \circ f \circ \sigma^{-1}(1) = \sigma(f(0)) = \sigma(\{0\}) = \{0\}$ , so  $f \neq \sigma^{(O \Rightarrow O)}(f)$ . The semantic value  $f$  fails to meet the invariance criterion because it distinguishes between different objects in the domain (objects and their singleton sets). The relation of set-theoretic inclusion fails to count as logical for similar reasons.

### 6.1.5 Advantages of this framework

By formulating the invariance criterion within a general semantic framework based on a typed presemantics, we give it a degree of generality lacking in most other formulations.<sup>14</sup> Instead of defining invariance for functions from one category to another, standard formulations talk either of extensions of predicates (construing quantifiers as second-level predicates: Tarski, Sher) or of sets of sequences that satisfy open formulas (construing quantifiers as connectives forming formulas from formulas: McGee, McCarthy). Both of these approaches produce formulations of the criterion that are less elegant and less general than the one presented above.

If we formulate permutation invariance in terms of *extensions*, then we must give separate clauses for the extensions of one- and two- place predicates, first- and second- level predicates, predicates with mixed argument places (i.e., with arguments in different types), and functors of various kinds. These proliferating clauses create unnecessary complexity. And yet the semantic values of sentential connectives do not fall under the analysis at all (Sher 1991:54).

If we take the other approach and frame the criterion in terms of satisfaction sequences, then we can accommodate sentential connectives and quantifiers at the same time, conceiving their semantic values as functions from sequences to sequences. But separate clauses are still required for connectives, predicates, and functors (McCarthy 1981, 510 n 12; McGee

<sup>14</sup>For example, Tarski 1966, McCarthy 1981:510, Sher 1991:55, Hodges in Sher 1996:678 n. 28, McGee 1996:570. An exception is van Benthem 1989, who works within a type theory and gives substantially the same definition as I have given above (318). (Where I define the induced transformation  $\sigma^{(X \Rightarrow Y)}(w)$  by means of composition and inverse,  $\sigma^Y \circ w \circ (\sigma^X)^{-1}$ , van Benthem defines it by specifying a set of ordered pairs,  $\{(\sigma^X(u), \sigma^Y(v)) : (u, v) \in w\}$ .)

1996 considers only connectives), and further tinkering is needed in order to accommodate higher types (McGee 574-5).

The approach presented above, by contrast, requires only one clause for the semantic values of connectives, predicates, functors, singular terms, and any other categories one dreams up, so long as they can be defined functionally in terms of the basic categories. It handles items in all (finite) types without extension, since semantic values for second-, third-, and higher-order quantifiers are already included in the type hierarchy (e.g., second-order quantifiers take values from  $((O \Rightarrow V) \Rightarrow V) \Rightarrow V$ ). It handles second-level predicates with mixed argument types equally simply: a predicate that takes an individual and a two-place relation as an argument, for example, would take a value from the type  $(O \Rightarrow ((O \Rightarrow (O \Rightarrow V)) \Rightarrow V))$ . Moreover, there is no need to provide explicitly for variable arities, since the semantic value of an  $n$ -place relation is a function from objects to semantic values for  $(n - 1)$ -place relations.

In addition to its elegance and generality, the articulation of permutation invariance presented here has some *philosophical* advantages. First, because it applies to types corresponding to *all* the grammatical categories, it avoids the narrow focus on sentential connectives and quantifiers characteristic of many discussions of logicality (see section 2.3.3, above). Second, it can be extended in a natural way to multivalued and intensional languages, as we shall see in sections 6.4 and 6.5, below. Most important, it makes explicit some assumptions that are hidden from view in standard presentations of the invariance criterion:

- that the permutations at issue are permutations of a *type* (O).
- that the other basic type (V) is *not* being permuted.

Both of these assumptions are concealed if we suppress reference to V in stating the permutation invariance criterion (as is common when invariance is applied to extensions or sets of satisfaction sequences) and talk about permutations of “individuals” without explicit relativization to a type. Making them explicit is the first step in the critique of the invariance

criterion that will be developed in sections 6.4–6.6.

## 6.2 Logical notions and logical constants

### 6.2.1 Presemantics and semantics

The permutation invariance criterion presented in section 6.1 is a criterion for the logicality of *semantic values*, not linguistic expressions. That is, it does not tell us directly which *expressions* in an (interpreted) language count as “logical constants;” rather, it tells us which *notions* (i.e., which items in each type) are logical. It makes no reference to specifically *semantic* relations between linguistic entities and objects or truth values: it defines logicality at the *presemantic* level.<sup>15</sup>

By “presemantics,” I understand (after Belnap) that part of semantic theory that makes no reference to linguistic expressions or their use. The basic task of presemantics is to give a theory of the semantic values that will be assigned to expressions in semantic interpretations. Just as it is often useful to separate grammar from semantics, so it is often useful to separate presemantics from semantics proper. Semantics proper brings together grammatical and presemantic concepts to give an account of how the semantic values of expressions depend on the semantic values of their parts.<sup>16</sup> Later (in section 6.7), it will be useful to distinguish semantics proper from *postsemantics*, in which the semantic values assigned to sentences are related to proprieties for the *use* of these sentences (generally by relating them to the top-level semantic notions of truth and implication). The relations between these disciplines are summed up in figure 6.1.

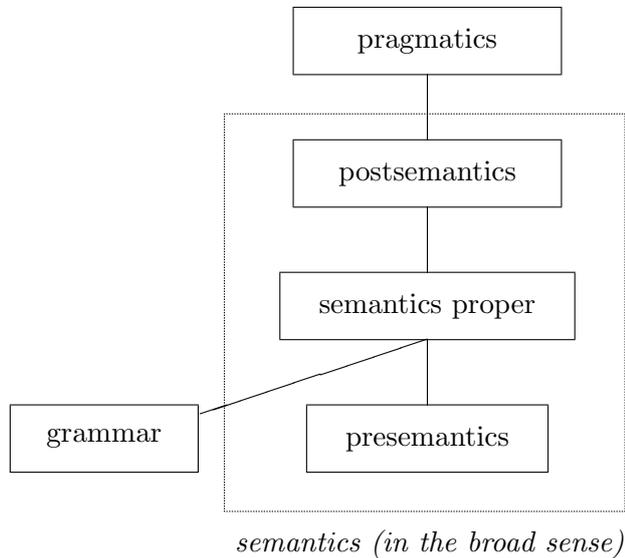
By articulating the invariance criterion at the presemantic level, then, we isolate logicality from *grammatical* considerations specific to particular notational systems.<sup>17</sup> I contend

<sup>15</sup>In this respect it follows Tarski 1966 and McGee 1996, whereas Sher 1991 operates at the semantic level.

<sup>16</sup>Grammar and presemantics are independent of semantics in the sense that they do not employ peculiarly semantic concepts. This is not to deny that grammatical and presemantic concepts are tailored to their use in semantics.

<sup>17</sup>Here I am indebted to Nuel Belnap’s unpublished manuscript UCL, §A.

Figure 6.1: Components of the semantic enterprise.



that this is an advantage in giving a criterion for logicality. The difference between logic and non-logic should not be sensitive to merely notational differences. If one gives a criterion for logicality at the semantic (as opposed to the presemantic) level, one runs the risk of entangling it with specific grammatical or notational features of the language for which one is doing semantics.

### 6.2.2 Logical notions and logical constants

Of course, the primary *application* for the concept of logicality is at the semantic level: defining implication and logical consequence. And for these purposes, it is not enough to have a criterion for logical *notions*: we will need to pick out a privileged class of linguistic *expressions*, the “logical constants.” However, the relationship between logical constants and logical notions is by no means straightforward. The obvious bridge principle,

An expression is a logical constant just in case it has a logical notion as its semantic value,

is clearly inadequate, because an intuitively non-logical expression may have a logical notion as its semantic value purely “by accident.” For example, suppose that the type  $O$  contains no two objects of the same mass. Then the two-place predicate “ $M$ ”, defined as holding between two objects just in case they have the same mass, has the identity relation (a logical notion) as its semantic value:  $M$  holds between  $x$  and  $y$  just in case  $x=y$ . But that does not make it a logical constant. Nor would it help to demand that logical constants *necessarily* have logical semantic values. For since water is necessarily  $H_2O$ , the two-place sentential connective  $H$ , defined as negation if water is  $H_2O$  and otherwise the constant-false truth-function, necessarily has a logical notion as its semantic value (the truth-function negation). Yet  $H$  is not intuitively a logical constant either.<sup>18</sup> It appears that whether a meaningful linguistic expression is a logical constant depends not only on its (actual or possible) semantic values, but on its meaning—that is, on the way its semantic value is determined. As McCarthy puts the point, “. . . the logical status of an expression is not settled by the functions it introduces, independently of how those functions are *specified*” (1981:516).<sup>19</sup>

Thus, in order to give necessary and sufficient conditions for an expression to be a logical constant, we would have to venture into the theory of meaning. I do not propose to do that here. Instead, I will abstract from considerations of the “rule” or “meaning” by means of which the semantic value of the expression is determined and consider only the way its semantic value varies with interpretations. That is, I will consider only the set of values  $\{i(t) : i \text{ is an interpretation}\}$  for a term  $t$ . Because I will not concern myself with the “semantic rule” by which these values are specified, I will not get a *sufficient* condition for logicality; but I will get a strong and interesting necessary condition.

This procedure may seem circular. After all, which semantic values are in  $\{i(t) : i \text{ is an interpretation}\}$  depends on which functions from expressions to semantic values count

<sup>18</sup>The example is due to McGee 1996:578. Those who don’t think that water is necessarily  $H_2O$  can use a necessarily true sentence of their choice, e.g. “ $2+2=4$ .”

<sup>19</sup>Objections of this general kind are raised by McCarthy 1981:514, McGee 1996:578, and Hanson 1997:391-2, and discussed by Peacocke 1976:228-9 and Sher 1991:64.

as *interpretations*. And whether such a function is an interpretation depends in part on which lexical primitives are logical constants, since these are constrained to receive certain semantic values. For example, there is no interpretation of standard first-order logic that assigns the constant “&” the semantic value of Boolean OR. Thus it looks as if I am assuming that it is determined what counts as an interpretation in order to determine which expressions are logical constants, when in reality the latter task presupposes the former.

But this circularity need not trouble us, if we understand the necessary conditions I will give as constraining both which terms are logical constants and which functions from expressions to semantic values are interpretations. This is reasonable, because the two questions are intertwined: the only reason we *care* which expressions are logical constants is that they are treated differently from non-logical constants in interpretations. Indeed, this differential treatment is what *makes* them logical constants. To say what a logical constant is, then, we need to say what it is for an interpretation function appropriately to treat a term  $t$  as a logical constant.

Without further ado, then, here are the conditions for logical constancy:

**Logical Constants:**  $t$  is a *logical constant* just in case

- **Simplicity:**  $t$  is a lexically primitive constant term. That is,  $t$  is not a variable and  $t$  is not composed of grammatically simpler expressions.
- **Logicity:**  $\{i(t) : i \text{ is an interpretation}\}$  contains only logical semantic values (as defined by the invariance criterion in section 6.1.4, above). That is, every interpretation assigns a logical semantic value to  $t$ .
- **Constancy:**  $\{i(t) : i \text{ is an interpretation}\}$  is a singleton set. That is, every interpretation assigns the same semantic value to  $t$ .

The three conditions deserve individual comment.

Simplicity is perhaps dispensable: the thought is that there is no *need* to define logical constancy for complex expressions, because their semantic values are determined recursively from the semantic values of their parts. It matters which primitive expressions are logical,

because this will affect the interpretation function; it does not matter in the same way which complex expressions are logical. But precisely because nothing turns on which complex expressions we call logical constants, one might do without Simplicity. One would then have to accept complex expressions like “ $\lambda x(x=x \vee Mx)$ ” as logical constants, even though they contain non-logical constants.

The point of Logicality is to exclude constants that respect individual characteristics of objects. Suppose the language contained a predicate “Red,” and interpretations were constrained to assign it extensions containing only red things. Then “Red” would not count as a logical constant, because it receives non-invariant semantic values.

Finally, the point of Constancy is to exclude constants that receive different logical semantic values in every interpretation, depending on which objects are in the domain. McGee’s charming example of such a constant is “wombat disjunction,” which acts like disjunction when the domain of quantification contains wombats and conjunction otherwise (1996:575). On any given domain, wombat disjunction receives a permutation-invariant semantic value (that of either disjunction or conjunction); yet it respects the individual characteristics of objects in just the way the permutation invariance account was intended to rule out.

Note again that these conditions cannot be sufficient for logical constancy, because they pay no heed to the rule by which the semantic values of an expression on an interpretation are determined. To return to McGee’s example, suppose that the semantic value of “H” on an interpretation  $i$  is given by the rule:

$i(\text{“H”}) = \text{Boolean NOT if water is H}_2\text{O, the constant FALSE truth function otherwise.}$

Since water *is* in fact  $\text{H}_2\text{O}$ ,  $i(\text{“H”}) = \text{Boolean NOT}$  on every interpretation. Since “H” is a lexical primitive and Boolean NOT is a logical semantic value, “H” meets the three conditions for logical constancy. Yet it is not logical.

For this reason, we cannot say that an expression is non-logical just in case it fails to meet the conditions for logical constancy. What we *can* do is lay down parallel necessary

conditions for non-logical constants:

**Non-logical Constants:**  $t$  is a *non-logical constant* just in case

- Simplicity:  $t$  is a lexically primitive constant term.
- Variability:  $\{i(t) : i \text{ is an interpretation}\}$  is the presemantic type corresponding to the grammatical category of  $t$ . That is, every value in the type gets assigned to  $t$  on *some* interpretation.

As before, Simplicity is included only because the distinction between logical and non-logical constants needs to be drawn only among lexically primitive expressions. The Variability condition is the complement of the Constancy condition for logical constants: it ensures that non-logical constants are genuinely “schematic.” Note that I have not built in an independence condition, which would require that every *combination* of assignments of values to expressions (respecting type) is realized as an interpretation. Thus, “cross-term” restrictions on interpretations like “the extension of ‘horse’ must be included in the extension of ‘mammal’” are not excluded.

We can now *postulate* that

**Postulate.** All lexically simple expressions are either logical or non-logical.

Together with the necessary conditions for logical and non-logical constants given above, this postulate imposes significant constraints on interpretation functions. It rules out any expression  $t$  such that  $\{i(t) : i \text{ is an interpretation function}\}$  is a proper subset of the type corresponding to the grammatical category of  $t$ , except where  $\{i(t) : i \text{ is an interpretation function}\}$  is a singleton set containing a logical semantic value.

The resulting picture is simple, attractive, and familiar. There are two types of primitive constant terms. The logical ones get the same semantic value in every interpretation; this value must be a logical (permutation-invariant) value. The non-logical ones get different values on different interpretations; their values can range through their whole semantic type.

### 6.2.3 Quantifiers and domains

Attractive as it is, this picture needs some modification. As it stands, it leaves no room for the standard quantifiers. The problem is that the semantic value of a quantifier in an interpretation depends on the *domain* parameter of the interpretation. For example, if the domain is  $\{1, 2, 3\}$ , then the universal quantifier is the function that takes just those functions in  $(O \Rightarrow V)$  that are true on 1, 2, and 3 to True. If the domain is  $\{3, 4\}$ , then the universal quantifier is a different function: the one that takes just those functions in  $(O \Rightarrow V)$  that are true on 3 and 4 to True. Thus the semantic values of the quantifiers vary with interpretations, just as the semantic value of wombat disjunction does, and the quantifiers fail to meet the Constancy condition.

They also fail to meet the Logicality condition. For the domain of an interpretation will in general be a proper subset of the type  $O$ . In that case, the semantic value of the quantifiers on that domain will not be invariant under *all* permutations of  $O$ , since some permutations of  $O$  will permute objects in the domain with objects outside of the domain. For example, the semantic value of the universal quantifier in an interpretation with domain  $\{1, 2\}$  is the function  $f$  that takes the value True on any function in  $(O \Rightarrow V)$  that takes 1 and 2 to True, and the value False for all other functions in  $(O \Rightarrow V)$ . Let  $\sigma$  be the permutation of  $O$  that switches 1 and 3 and takes all other objects to themselves. Let  $g$  be a function that takes the value True on 1 and 2 and False on 3. Then  $f(g) = \text{True}$ . But  $\sigma^{((O \Rightarrow V) \Rightarrow V)} f(g) = f \circ (\sigma^{(O \Rightarrow V)})^{-1}(g) = f((\sigma^{(O \Rightarrow V)})^{-1}(g)) = f(g \circ \sigma^{-1}) = \text{False}$  (because  $g \circ \sigma^{-1}$  takes 1 to False). So  $f$  is not invariant under the transformation induced by  $\sigma$ : that is,  $f$ , the semantic value of the universal quantifier in the domain  $\{1, 2\}$ , is not a logical semantic value.

Thus the universal quantifier fails to meet the necessary conditions laid down in section 6.2.2 for logical constancy. But it does not meet the conditions for non-logical constancy, either, and thus falls afoul of the postulate that every lexically primitive constant is either logical or non-logical. For the set  $\{i(\text{“}\forall\text{”} : i \text{ is an interpretation})\}$  falls far short of the type  $((O \Rightarrow V) \Rightarrow V)$ , and “ $\forall$ ” fails to meet the Variability condition.

This problem could be solved by changing the criterion for logical semantic values. In-

stead of demanding invariance under all permutations of  $O$ , we could demand invariance under all permutations of the *domain* (this is essentially what Sher 1991 does). But this approach amounts to abandoning the attempt to define logicity at the presemantic level, since the invariance criterion will now have to be relativized to domains, a feature of (properly semantic) interpretations. We will no longer have a criterion for logicity of *semantic values*, but rather for the logicity of a semantic value relative to an interpretation or domain.

Moreover, this approach leaves the problem of Constancy untouched: quantifiers will still receive different semantic values in different interpretations. Sher 1991 deals with this problem by getting rid of the Constancy condition (48). She rules out constants that behave like wombat disjunction (in that they are sensitive to the particular identities of objects in the domain) in a different way—by requiring that logical constants be invariant under arbitrary *bijections* from one domain to another (of the same cardinality). Wombat disjunction is disqualified because its semantic value is not invariant under a bijection that maps a domain containing wombats to a domain not containing wombats. This solution is adequate, but it takes the invariance criterion even further from the presemantic level.

I am going to take a different approach to the problem of quantifier domains, one that solves both problems (Logicity and Constancy) in one stroke and allows us to keep the invariance criterion at the *presemantic* level. The idea is to make the presemantic types *indexical*, that is, context-relative. The quantifier domain can then be understood as the value of the type  $O$  at a context  $c$ ,  $O(c)$ , and the semantic value of a quantifier as a function from contexts  $c$  to items in the function space  $(O(c) \Rightarrow V) \Rightarrow V$ . This approach (which will be motivated and explained in more detail in section 6.3, below) allows quantifiers to meet the Constancy condition, because in every interpretation they are assigned the same function from contexts to functions. It also allows them to meet the Logicity condition, when the invariance criterion is appropriately retooled to apply to indexical presemantic types (section 6.3.4, below). Thus logicity can still be defined at the presemantic level—at the cost of complicating the presemantics somewhat.

### 6.2.4 A note on the status of $\lambda$

Before we move on to the details of this approach, it is worth commenting briefly on the status of the  $\lambda$  operator. Is it a logical constant? It is certainly not a variable or a non-logical constant. But it does not meet the criteria given above for logical constants, either, since its meaning is not given by a function that assigns it a semantic value in every interpretation. Yet it will be an essential component of any logically true sentence involving variable binding. Taking seriously Cresswell's (1973) claim that  $\lambda$  "...is perhaps not the usual sort of logical constant like the PC truth functors and may well be a constant which acts at a rather deeper level than they do" (88), I will call it a *structural constant*. In giving it this name, I mean to class it with the *structural features* of sentences that bear on logical truth: the order, arrangement, and grammatical categories of symbols. It is these that are left when we abstract from all of the expressions that can take on a semantic value in an interpretation. The grammar described in section 6.1.1 does not have any structural constants besides  $\lambda$ , but the parentheses used in many formal languages are structural constants.<sup>20</sup>

If one wanted to give a sense to *3-formality* in the present framework, it would be natural to invoke these structural features. An 3-formal logic would not have any "logical constants": it would abstract *entirely* from content, even from the abstract and general content possessed by logical terms. It would consider only the semantic properties of sentences that depend on *structural features*. There would be no 3-formal truths, but there might be non-trivial 3-formal consequences. In a language with with the  $\lambda$  operator, for example, consequences licensed by the rules of  $\lambda$ -conversion would be 3-formal.<sup>21</sup> Evans 1985 suggests that some inferences in English may be valid solely on the basis of the order, arrangement, and semantic

<sup>20</sup>"Syncategorematic terms" is an older name for structural constants. Note that although the logical constants are often taken to be syncategorematic, they are not syncategorematic in the present framework, which tries to preserve as many semantic and grammatical parallels as possible between logical and non-logical terms.

<sup>21</sup>These rules allow one to intersubstitute formulas of the forms ' $(\lambda x\phi x)a$ ' and ' $\phi a$ ', provided the categories of ' $\phi$ ', ' $x$ ', and ' $a$ ' are appropriate, and to reletter variables. See Church 1940:60 and Cresswell 1973:88.

categories of expressions: for instance, the inference from “Dan is a large man” to “Dan is a man” (54). I will not discuss this notion of “structurally valid” inferences further in what follows, but it is worth considering whether it is a reasonable surrogate for the notion of 3-formal inference. If so, then only *part* of logic is 3-formal.

## 6.3 Invariance over variable domains

### 6.3.1 What does the quantifier domain represent?

The root of the problem with the quantifiers (raised in section 6.2.2, above) is the use of different quantifier domains in different interpretations or models. If we used the same domain in every interpretation, we could assign to each quantifier a fixed, permutation-invariant semantic value in the type  $((O \Rightarrow V) \Rightarrow V)$ . The use of variable domains, on the other hand, seems to require that quantifiers receive different semantic values—not all of them permutation-invariant—in interpretations with different domains.

Thus it is worth asking: why do we use variable domains in logical semantics? More broadly, what is represented by the domain parameter of an interpretation? When is it appropriate to use one domain rather than another? In order to answer these questions, we must consider more generally what is represented by a semantic interpretation or model.<sup>22</sup>

Etchemendy 1990 articulates a common view of the options: either models represent the semantic values the (fully interpreted) simple terms of the language would have in various possible states of the world, or they represent the semantic values the simple terms would have in the actual world if their meanings were different (respecting grammatical categories). In short, semantic interpretations represent either possible worlds or possible meanings. Each of these ways of thinking about interpretations suggests a different view of the domain parameter: neither, however, is fully satisfactory.

On the “possible worlds” approach, the domain of an interpretation represents the col-

---

<sup>22</sup>I will use these terms interchangeably. Often a model is articulated into a domain and an interpretation function: “interpretation” as I use it encompasses both components.

lection of objects that exist at some possible world.<sup>23</sup> But this does not seem right, for several reasons.

First, it does not fit our standard *use* of domains. Our intended interpretations of theories often employ domains containing only a subset of the objects we think exist in the actual world. When we take the intended domain of a theory to be, say, the set of integers, we are *not* engaging in counterfactual reasoning and asking what would be true in a world where only the integers exist: we are reasoning about the integers in the actual world.

Second, if domains represent the objects existing in different possible worlds, then our standard practice of allowing any set to count as the domain of an interpretation (e.g., in assessing validity) depends on the assumption that given any set of objects, the world could have contained just those objects. This is a substantial and contentious assumption on which to base one's logical theory: it implies, for instance, that mathematical objects are not necessary existents.<sup>24</sup>

Third, this approach does not make good sense of the analogous use, in modal logic, of different *frames* or *model structures* (collections of possible worlds with accessibility relations) in different interpretations. Perhaps one could make sense of the idea that modal frames represent different ways the structure of possibility *might have been*—different possible “superworlds” filled with possible worlds<sup>25</sup>—but the idea does not initially seem promising.<sup>26</sup> It would be preferable to have an account of quantifier domains that would help to

<sup>23</sup>For an influential example of this approach, see Kaplan 1989:543-5.

<sup>24</sup>See Hanson 1988b:387. This objection is sometimes met by claiming that the sense of possibility at issue is only *logical* possibility (e.g., Shapiro 1998:147-8). Hanson's own response to the objection is to construe the domain of an interpretation as the set of objects in a “subworld” of a possible world.

<sup>25</sup>For example, we might take the frames to represent different *epistemically* possible frameworks of *metaphysically* possible worlds.

<sup>26</sup>Cf. Pollock 1967: “... why are we allowed to construct models having any number of possible worlds? After all, there is only one set of all possible worlds, and that has a unique cardinality. How are all of these different sets of assignments, of all different cardinalities, supposed to each represent the set of all possible worlds?” (134-5). Hanson and Hawthorne 1985: “Are we to think of the model structures other than S that are involved in the definition of validity<sub>2</sub> as representing ways that possible worlds might have been? Surely this is unsatisfactory. The whole point of invoking possible worlds in the first place was to provide a foundation for our notion of possibility. Any attempt to analyze this notion in terms of possible sets of possible worlds would be gratuitous” (15).

illuminate the formally analogous use of modal frames.

Does the “possible meanings” approach yield a better understanding of domains? On this approach, quantifiers are construed not as having fixed meanings (“everything,” “something”), but as being given their meanings by interpretations, just like non-logical terms. The domain varies in response to variation in the meaning of the quantifiers. Thus, an interpretation whose domain is the set of dogs can be thought of as assigning to “ $\forall$ ” the meaning “all dogs” and to “ $\exists$ ” the meaning “some dogs.”<sup>27</sup>

This approach makes better sense of our practice of using interpretations with restricted domains than the “possible worlds” approach does: when we take the domain of a theory to be the integers, we are simply reading “ $\forall x$ ” as “for all integers.” But as Etchemendy 1990 has argued, this approach fails to make sense of other features of standard semantic theory: the requirements that there be a single domain for the universal and existential quantifiers, that the semantic values for singular terms be drawn from the quantifier domain, that the semantic values for functors be functions defined over the quantifier domain, and so on.<sup>28</sup> For example, an interpretation whose domain is the set of dogs cannot make Nixon (the man) the semantic value of the singular term “Nixon.” Etchemendy 1990 calls these requirements “cross-term restrictions,” because they restrict the semantic values that can be assigned to one term on the basis of the semantic value assigned to a different term (68). But if interpretations represent possible meaning assignments, then these cross-term restrictions seem unmotivated. Why *shouldn't* we allow interpretations that take the existential quantifier to mean *some dogs* and the term “Nixon” to mean Nixon? We can give an *extrinsic* rationale for excluding such interpretations: they would admit counterexamples to inferences we regard as logically valid, such as “ $F_n \supset (\exists x)Fx$ .” But our construal of interpretations as possible meaning assignments does not *itself* justify the restrictions:

---

<sup>27</sup>See Etchemendy 1990:67. If we want to distinguish meaning from semantic value, we can say: the interpretation assigns to “ $\forall$ ” the semantic value it would have (in the actual world) had it meant “all dogs.”

<sup>28</sup>The reason for these requirements is clear. Without them, there would be counterexamples to inferences such as “ $Fa$ , therefore  $(\exists x)Fx$ ” (when the semantic value of “ $a$ ” is outside the quantifier domain) and “ $(\forall x)(Fx \supset Gx)$ ,  $(\exists x)Fx$ , therefore  $(\exists x)Gx$ ” (when the domain of  $\exists$  is wider than the domain of  $\forall$ ).

there is nothing *incoherent* about simultaneously interpreting “ $(\exists x)$ ” to mean *some dogs* and “Nixon” to mean Nixon. To the extent that we would like semantic theory to be *explanatory* of validity and invalidity, this is a serious deficiency.

The “possible worlds” construal of interpretations makes better sense of cross-term restrictions. If the domain is just the set of all the things that exist (at a world), then it is reasonable to limit the referents of names to that set (and similarly for the semantic values of functors). But as we have seen, this construal makes no sense of our use of restricted domains for reasoning about the actual world. Nor does it help to combine the two approaches, taking domains to be the semantic values of quantifiers with some possible meaning in some possible world. For since this solution permits the quantifier domain to be a proper subset of the set of objects, the problem of motivating cross-term restrictions arises for it, too.

In sum, neither the “possible worlds” nor the “possible meanings” construal of interpretations makes sense of the way interpretations are used in standard semantic practice: the former cannot account for the use of restricted domains, the latter for cross-term restrictions.

### 6.3.2 Quantifier domains as indexical types

I want to suggest that the domain of an interpretation represents neither the collection of objects existing at some possible world, nor the meaning of some possible quantifier term, nor a combination of the two (the meaning of some quantifier at some world). Rather, the domain is a specification of a presemantic *type*: the type O, from which semantic values for singular terms are to be drawn in an interpretation.

This account makes beautiful sense of the “cross-term restrictions” in standard semantics. The reason that semantic values for singular terms cannot come from outside the quantifier domain is that the quantifier domain is just the presemantic type O. What now needs explaining is how different domains can be used in different interpretations. On this account, it seems, the use of variable domains must be construed as the use of *different pre-*

*semantic types* to supply semantic values for expressions of the *same grammatical category* on different interpretations. On one interpretation, we take semantic values for singular terms from  $O_1$ , on another, from  $O_2$ . And surely this consequence is much more troubling than the problem of cross-term restrictions! We had taken for granted as a fundamental constraint on interpretations that semantic values for terms of a given grammatical category be taken from the correlated presemantic type. If we now relax this constraint, on what grounds can we exclude interpretations that assign a unary truth function to “F” and a truth value to “x” in “Fx”?

The solution I propose is to take  $O$  to be an *indexical* presemantic type. By an indexical type, I mean a type that provides a definite range of values only relative to an index. The index to which  $O$  is relative is a *sortal* concept (or a collection of sortal concepts, together with stipulations for assessing identity claims between objects picked out under different sorts) that specifies what counts as an “object” on an interpretation. In suggesting that the type  $O$  is relative to a sortal concept or concepts, I am only reiterating the familiar point that “object” and “thing” are not themselves sortal concepts, but rather *prosortals* that stand in for contextually understood sortals (cf. Brandom 1994:438). When we say “there is nothing on Joe’s desk,” we are not saying that there are no specks of dust, air molecules, shadows, or two-dimensional regions, but merely that there are no medium-sized material objects (pencils, papers, etc.). Similarly, when we say that nothing satisfies an equation, we may mean only that no *real* number satisfies it, or that no *integer* does. What counts as a *thing* in some stretch of discourse is determined largely by context and conversational convention.<sup>29</sup> In this respect, the type  $O$  resembles indexical expressions: just as “this” and “I” receive a determinate semantic value only relative to some context, so, I suggest,  $O$  provides a determinate *range* of semantic values only relative to a context.<sup>30</sup> Mathematically, we can think of  $O$  as a function from contexts to sets of values.

<sup>29</sup>Cf. Stalnaker 1970:35.

<sup>30</sup>Cf. Wright 1998: “The contribution of its range of quantification to the purport of a quantified statement is best compared to the contribution effected by context. . . to the purport of a statement involving *indexicals*” (345-5).

In fact, we get a simpler theory by thinking of *all* the simple types as indexical. The types that do not change with context (such as  $V$  in standard two-valued logic) can be represented as *constant* functions from contexts to sets of values. In some semantic frameworks, however,  $V$  exhibits the same kind of indexicality as  $O$  does in extensional quantifier logic. For example, consider a modal sentential logic in which the semantic values for sentences are taken from the type  $V$  of *propositions*, conceived as functions from possible worlds to truth values. The set of possible worlds varies from one interpretation to another—it is specified in the modal *frame*—and thus so does the available stock of propositions. In such semantic frameworks,  $V$  can be construed as an indexical type (indexed to a contextually and conventionally determined set of possible worlds).

### 6.3.3 Semantics with indexical types

The provision for indexical expressions and indexical types complicates the semantics somewhat.

At the presemantic level, we must add the notion of a *context* and replace the old notion of “semantic value” with two new notions, *content* and *character*.<sup>31</sup> A *content* is a “determinate” semantic value, the kind an expression has at a given context: for instance, an object, a truth value, or a function from objects to truth values. A *character* is an “indeterminate” or context-relative semantic value, that is, a *function* from contexts to contents. Thus, on the intended interpretation of English, the word “I” has the same character for all speakers, but its content depends on an aspect of the context of utterance: it always denotes the speaker. Since both contexts and contents are presemantic, so are characters.

*Types* can now be construed as functions from contexts to sets of contents. For example,  $O$  can be construed as a function from contexts to sets of objects (the domains at those contexts). Although types are not sets of characters, we can say that a character  $x$  is “in” a

<sup>31</sup>The terms come from Kaplan 1989, though my semantic framework is different from Kaplan’s in several respects.

type  $t$  just in case for every context  $c$ ,  $x(c)$  is a member of  $t(c)$ , the set of contents allowed at  $c$  for type  $t$ .

At the semantic level, we now take an *interpretation* to be an assignment of a *character* (not a content) in the appropriate type to each expression of the language. Given the characters of the simple expressions of a language, we can calculate the characters of complex expressions composed from them using a recursive definition. Note that *interpretations are context-independent*: the interpretation of a language does not vary with changes in the context. “I” can have the same interpretation in my mouth as it has in yours, though it receives a different content.

Within this semantic framework, we can assign *all* logical constants (including quantifiers) a constant character across interpretations. The sensitivity of the quantifiers to the domain of quantification is now just a special case of the dependence of content on context that is captured in the general notion of character. The same character can determine a different content on different domains, because the domain is one component of context. Thus we can assign a fixed character to each logical constant, including the quantifiers, and require that an interpretation assign each logical constant its proper character. What is particularly nice about this strategy is that it addresses the problem of the apparent non-constancy of the quantifiers in just the same way as it addresses the problem of the apparent non-constancy of “I” and “this”: by distinguishing between a constant semantic value (character) and a non-constant, context-sensitive one (context). This is a distinction we need to make anyway if we are to accommodate indexicals; it is not just an ad-hoc move designed to patch up the conditions for logical constancy given in section 6.2.2.

Finally, at the postsemantic level, we give an account of the conditions under which sentences are *true* (given their characters) and of when one sentence is a *logical consequence* of a set of other sentences (given their characters). The truth of a sentence will, in general, be relative to an interpretation and a context of utterance. Consequence, by contrast, is independent of both interpretation and context, because it is defined in terms of a quantification over all contexts and all logical interpretations. (Note that the usual quantification

over all *domains* in the definition of consequence is included in the quantification over all contexts.)

### 6.3.4 The invariance criterion restated

We can now reformulate the invariance criterion as a criterion for the logicality of *characters*. The simplest way to do this would be to call a character logical just in case it determines a permutation-invariant content at every context (where permutation invariance for contents is defined as in section 6.1.4). But this definition does not exclude the character of “wombat disjunction” (see section 6.2.2, above), which yields different contents depending on whether or not the domain of the context contains wombats. A character with this kind of sensitivity to the particular differences between different objects should not count as logical. To solve this problem, we will demand not only that the contents determined by logical characters at each context be permutation-invariant, but that they be invariant under arbitrary *bijections* from one domain into another of the same cardinality.<sup>32</sup>

As before, we will need to define how a bijection from  $O(c)$  to  $O(d)$ , for arbitrary contexts  $c$  and  $d$ , induces a bijection from  $Z(c)$  into  $Z(d)$  for all derived types  $Z$ :

**Induced Transformation.** Where  $c$  and  $d$  are contexts and  $\sigma$  is a bijection from  $O(c)$  to  $O(d)$ ,  $\sigma^Z$  is the *bijection from  $Z(c)$  into  $Z(d)$  induced by  $\sigma$* . It is defined inductively on the complexity of  $Z$ :

if  $Z=O$ , then for all  $w \in Z(c)$ ,  $\sigma^Z(w) = \sigma(w)$

if  $Z=V$ , then for all  $w \in Z(c)$ ,  $\sigma^Z(w) = w$

if  $Z=(X \Rightarrow Y)$ , for any types  $X, Y$ , then

for all  $w \in Z(c)$ ,  $\sigma^Z(w) = \sigma^Y \circ w \circ (\sigma^X)^{-1}$ .

---

<sup>32</sup>Cf. McGee 1996:576: “. . . a logical operation across domains ought to be left fixed by an arbitrary bijection. A property which, while invariant under all permutation [sic] of a given domain, is disrupted when we move, via a bijection, to a different domain must depend on some special feature shared by the members of the first domain and lacked by the members of the second domain. It is not the sort of purely structural property that pure logic studies.” Here McGee follows Sher 1991, except that McGee is operating at a presemantic level (with “operations across domains”), Sher at a semantic one (with “logical constants” and their interpretations in different models).

We can now define bijection invariance as follows:

**Bijection Invariance.** A character  $w$  in type  $Z$  is *bijection invariant* iff

for every context  $c$ ,

for every context  $d$  such that  $O(d)$  has the same cardinality as  $O(c)$ ,

for every bijection  $\sigma$  from  $O(c)$  to  $O(d)$ ,

$\sigma^Z(w(c)) = w(d)$ .

From this we get the following criterion for the logicality of characters:

**Invariance Criterion.** A character  $w$  in type  $Z$  is *logical* iff it is bijection invariant.

The characters proper to the existential and universal quantifiers—indeed, all the cardinality quantifiers—come out logical on this criterion.

### 6.3.5 Logical constants revisited

We have now removed the obstacle to accepting the necessary conditions for logical constancy proposed in section 6.2.2, according to which logical constants are expressions that receive the same (invariant) semantic value in every interpretation. The quantifiers appeared to be an exception, since they receive different semantic values (in the sense of *content*) on different interpretations, and not all of these values are invariant. But we can solve this problem by thinking of the type  $O$  as indexical, as I have suggested, and remembering the bifurcation of “semantic value” into content and character. Although the quantifiers receive different *contents* in different contexts and on different interpretations in the same context, they have the same *character* on every interpretation. Moreover, it is an invariant character, as defined above.

Thus, all that is necessary to salvage the necessary conditions for logical and non-logical constants is to replace references to semantic value with references to character:

**Logical Constants:**  $t$  is a *logical constant* just in case

- **Simplicity:**  $t$  is a lexically primitive constant term. That is,  $t$  is not a variable and  $t$  is not composed of grammatically simpler expressions.
- **Logicality:**  $\{i(t) : i \text{ is an interpretation}\}$  contains only logical characters (as defined by the bijection invariance criterion in section 6.3.4, above). That is, every interpretation assigns a logical character to  $t$ .
- **Constancy:**  $\{i(t) : i \text{ is an interpretation}\}$  is a singleton set. That is, every interpretation assigns the same character to  $t$ .

**Non-logical Constants:**  $t$  is a *non-logical constant* just in case

- **Simplicity:**  $t$  is a lexically primitive constant term.
- **Variability:**  $\{i(t) : i \text{ is an interpretation}\}$  contains every character “in” the presemantic type corresponding to the grammatical category of  $t$ . That is, every character “in” the type gets assigned to  $t$  on *some* interpretation. (Recall that a character  $x$  is “in” a type  $t$  just in case for every context  $c$ ,  $x(c)$  is a member of  $t(c)$ , the set of contents allowed at  $c$  for type  $t$ .)

Note that nothing in the revised definition of Constancy rules out a logical constant whose content varies wildly with the context, provided its character is bijection-invariant. One could have a logical constant that acted like Boolean OR on domains with odd finite cardinality, Boolean NAND with domains with even finite cardinality, and Boolean AND on domains with infinite cardinality.

### 6.3.6 Advantages of this approach

By taking the presemantic type  $O$  to be indexical, we have solved two different problems. First, we have made good sense of the use (in formal semantics) of different domains in different interpretations. On the present account, the domain of an interpretation represents neither the ontology of a possible world nor a possible meaning of a quantifier term (or a component thereof). It is not a sub-range of the type  $O$ , but the type  $O$  itself (at a context).

This construal of the domain makes beautiful sense of the fact that the domain constrains assignments of semantic values not only to quantifiers, but to singular terms, predicates, and functors. These constraints need not be taken as *ad hoc* “cross-term restrictions,” as they must be in the “possible meanings” construal of the domain. The present proposal also makes sense of the use of domains smaller than the ontology of the actual world in non-counterfactual reasoning—a practice that is unexplained by the “possible worlds” construal of the domain.

Second, taking domains as specifications of (indexical) types allows us to think of all logical constants—quantifiers included—as having constant semantic values (in the sense of character, not content) across all (logical) interpretations. This allows us to restrict the application of the invariance criterion to the *presemantic* level. We need not refer to interpretations or models in defining invariance (as Sher 1991 does): it is *characters*, not logical constants, that are invariant or not. Logical constants can then be taken to be expressions that are assigned an invariant character by an constant extensional function from interpretations to characters.<sup>33</sup>

The point of this discussion was to show that the presemantic approach to invariance and the logical constants developed in sections 6.1 and 6.2 can be extended to the quantifiers. In the remaining sections, we will work in the simpler framework of section 6.1 (without indexical expressions or indexical types), so as not to clutter the philosophical points with unnecessary technical complexity. In this simpler framework, we can safely talk of the permutation invariance of semantic values (in the sense of section 6.1.4) rather than the more complex notion of bijection invariance of characters (in the sense of section 6.3.4). The material in this section should suffice to show how the points I make below about permutation invariance can be applied to the more complex case of bijection invariance on indexical types.

---

<sup>33</sup>Recall that this is only a necessary condition for logical constancy, not a sufficient one: see section 6.2.2, above.

## 6.4 Extending the account to sentential operators

One advantage of the way we have presented the invariance criterion (in sections 6.1.4 and 6.3.4, above) is that it makes it very easy to see an asymmetry in the treatment of the two basic semantic types,  $O$  and  $V$ . The only permutations considered are permutations of  $O$ . We do not demand that logical notions be invariant under permutations of  $V$ . The criterion works in the special case of two-valued logics only because in such logics (where  $V$  is the set  $\{\text{True}, \text{False}\}$ ), *every* operator in the categories  $(V \Rightarrow V)$ ,  $(V \Rightarrow (V \Rightarrow V))$ , etc., is logical.

This asymmetry is concealed in non-type-theoretic presentations of the invariance criterion (such as Tarski's, Sher's, McCarthy's, and McGee's). Instead of taking the semantic values of quantifiers to be items in the function space  $((O \Rightarrow V) \Rightarrow V)$ , these presentations take them to be sets of sets, or alternatively sets of satisfaction sequences—with no explicit reference to  $V$  or the truth values. Although this is only a difference in representation—it reveals no real disagreement about the *semantics* of the quantifiers—it is a significant one. If we mention  $O$  in our account of permutation invariance but leave  $V$  implicit, we conceal the asymmetry in their treatment. By making both types explicit and making the asymmetry manifest, we provoke the question of its justification. If we have reason to demand that logical notions be invariant under permutations of  $O$ —that is, insensitive to the differences between particular objects—why shouldn't we also demand that they be invariant under permutations of  $V$ —that is, insensitive to the differences between particular truth values?<sup>34</sup>

It is not a sufficient answer to note that if we did this, some paradigm logical notions—conjunction, disjunction, and the material conditional—would not qualify as logical.<sup>35</sup> All this shows is that we had better *find* a justification for treating  $V$  and  $O$  differently: it is not itself such a justification (or at any rate, only a very weak one). The invariance criterion

<sup>34</sup>Although van Benthem 1989 works in a type-theoretic framework, which forces him to *acknowledge* the assumption that “truth values retain their individuality” (318), he does not pursue the question of the justification of this assumption or its coherence with the general motivations behind the invariance criterion.

<sup>35</sup>Only four binary sentential operators are invariant under all permutations of (two-valued)  $V$ : the degenerate truth functions  $*$  such that  $A*B$  is equivalent to  $A$ ,  $B$ , the negation of  $A$ , and the negation of  $B$ , respectively.

is intended to be something more than just a curve fitting our antecedent intuitions about which notions are logical: it is supposed to help us see *why* we count these notions as logical, by embedding our intuitions about logicality in a broader theoretical framework. To the extent that it cannot explain the asymmetry between O and V, except by adverting to the disastrous consequences of treating them symmetrically, it fails in this endeavor. Absent some *principled* reason for demanding invariance under permutations of O but not permutations of V, we cannot claim that the invariance criterion explains the logicality of the truth-functional operators.

It might be objected that the motivations behind the permutation invariance criterion apply only to O. The point of the criterion is to ensure that logical notions do not respect particular features of *objects*. The issue of invariance under permutations of the *truth values*, it might be said, is simply irrelevant.

But not all of the particularity from which logic must abstract need be particularity of *objects*. This becomes clear when we consider non-classical logics, in which the type V is larger than the set {True, False}. Suppose, for example, that the type V is a set of *propositions* (in some sense), rather than truth values. This type might be appropriate for a language containing non-truth-functional sentential operators (regardless of whether any of these operators are logical constants). If we now apply the permutation invariance criterion as stated in section 6.1.4—permuting O but not V—then *every* item in the derived types  $(V \Rightarrow V)$  and  $(V \Rightarrow (V \Rightarrow V))$  will be counted logical. But most of the items in these types are clearly not logical notions: for example, the function in  $(V \Rightarrow V)$  that takes the value  $2+2=4$  on the argument *there are ocelots in Texas* and the value  $2+2=5$  on all other arguments, or the function that takes the value *Clinton knows that p* on the argument *p* (for any proposition *p*). These manifestly non-logical notions will not be excluded by the invariance criterion unless we consider permutations of V as well as O.

Thus, we have two options. We can restrict permutations to the type O and admit that the invariance criterion provides at best a necessary (but by no means sufficient) condition for a notion to be logical. Or we can require that logical notions be invariant not just under

permutations of  $O$ , but also under at least *some* permutations of  $V$ . What is tricky, if we take this second path, is specifying the *subclass* of permutations of  $V$  under which logical notions must be invariant. As we have seen, *all of them* is too many, and *none of them* too few. Can we give a principled justification for excluding some (but perhaps not all) permutations of  $V$  in applying the invariance criterion? Call this the *justificatory challenge*.

A natural way to do this would be to identify some *structure* on  $V$  and consider only permutations that preserve this structure: that is, *automorphisms* of the structure. One obvious candidate for such structure would be the division of propositions into true and false. If we call items in  $(V \Rightarrow (V \Rightarrow V))$  “logical” just in case they are invariant under every permutation of  $V$  that respects the division of propositions into true and false (that is, does not map a true proposition onto a false one or a false one onto a true one), then the logical notions in that type turn out to be just the binary truth functions and the identity relation on propositions.<sup>36</sup> As it stands, this criterion excludes modal, temporal, and other intensional operators: but these too might count as invariant if we required permutations to preserve *more* structure on  $V$  than merely the distinction between true and false.<sup>37</sup>

Accordingly, let us think of the basic presemantic types (e.g.,  $O$  and  $V$ ) as *structured* sets of entities or values rather than unstructured collections. Call the structure that is intrinsic to the type itself *intrinsic structure*, to distinguish it from the countless other structures that might be defined on the items in the type. Some types—for instance,  $O$ , as usually conceived—appear to have no intrinsic structure; but for uniformity of treatment, we will say instead that their intrinsic structure is the *null structure*. We can then reformulate the invariance criterion in a way that treats all the basic types symmetrically:

<sup>36</sup>Although Sher’s version of the invariance criterion does not apply to sentential operators, she hints that they can be thought of as “formal” on the grounds that they are invariant under all permutations that respect truth values: “When it comes to sentential connectives, we can regard their formality as based on ‘not distinguishing the identity of propositions.’ Intuitively, sentential connectives are formal iff they distinguish only patterns of propositions possessing truth values and nothing else. The interpretations of logical connectives as (denoting) Boolean truth functions reflects just this intuition” (1996:678).

<sup>37</sup>Intensional operators will be discussed further in the next section. Here I am trying to make a more general point.

**Induced Transformation.** Where  $\sigma^A, \sigma^B, \sigma^C$ , etc. are permutations of basic types A, B, C, etc.,  $\sigma^Z$  is the *transformation on the type Z induced by*  $\sigma^A, \sigma^B, \sigma^C$ , etc. It is defined inductively on the complexity of Z:

if Z is a basic type T, then for all  $w \in Z$ ,  $\sigma^Z(w) = \sigma^T(w)$   
 if  $Z=(X \Rightarrow Y)$ , for any types X, Y, then  
 for all  $w \in Z$ ,  $\sigma^Z(w) = \sigma^Y \circ w \circ (\sigma^X)^{-1}$ .

**Permutation Invariance.** A semantic value  $w$  in type Z is *permutation-invariant* iff for all *intrinsic structure-preserving* permutations  $\sigma^A, \sigma^B, \sigma^C$ , etc., of the basic types A, B, C, etc.,  $\sigma^Z(w) = w$ .

**Invariance Criterion.** A semantic value  $w$  in type Z is *logical* iff it is permutation-invariant.

Not only does this formulation of the invariance criterion treat all the basic types symmetrically, it does not even *mention* any particular basic type (e.g., O or V). Thus, it is not limited in its application to semantic values for extensional or two-valued languages. We will see in the next section how it can be applied to a presemantics with *three* basic types.

It should not be thought that by limiting ourselves to permutations that preserve a basic type's intrinsic structure, we have met the justificatory challenge (i.e., by what principle do we allow *some* permutations on V but not others?). We have only relabelled the problem. For now we need to know what is meant by intrinsic structure. Still, this relabelling is not a *mere* relabelling, for three reasons. First, in place of a problem about V, we now have a *general* problem about all of the basic types, including O. Second, by connecting the restrictions on permutations with the general notion of intrinsic structure, we avoid the need to encumber our formulation of the invariance criterion itself with details about particular types. Finally, when we do turn our attention on V, there are many cases in which there is an obvious *prima facie* candidate for intrinsic structure. By using these cases as a guide, we may be able to articulate a theoretical conception of intrinsic structure with more general applicability (this task will concern us in section 6.7).

The cases I have in mind are the presemantic frameworks appropriate for multivalued sentential logics. In these cases,  $V$  is a structured set of multivalues. The structure is that required for the definition of implication. In classical logic, a sentence  $R$  implies a sentence  $S$  just in case every interpretation that makes  $R$  true makes  $S$  true. Later generalizations of this idea for multivalued logics divide multivalues into *designated* and *undesignated* values and invoke *designatedness* instead of truth. But even this generalization is a special case of the most general approach: the imposition of a relation  $\leq$  on the values.<sup>38</sup> A sentence  $R$  implies a sentence  $S$ , on this account, just in case every interpretation gives  $R$  a value that bears  $\leq$  to the value it gives to  $S$ .<sup>39</sup> I suggest that the structure imposed on a set of multivalues by the  $\leq$  relation is a plausible candidate for “intrinsic structure” on a type  $V$ . In section 6.7, I will provide a theoretical basis for this claim; here, I simply want to explore its consequences in three different cases.

First, consider standard two-valued logic. There are two values, True and False, such that  $\text{False} \leq \text{True}$  and it is not the case that  $\text{True} \leq \text{False}$ . If we take  $\leq$  to be intrinsic structure, then there is evidently only one intrinsic structure-preserving permutation: the trivial (identity) permutation. Hence all of the two-valued truth functions come out as logical on the modified invariance criterion.

Second, consider Anderson and Belnap 1992’s four-valued logic for tautological entailment. In this logic, the values represent epistemic states of a computer database with respect to particular propositions: “told only true” (**t**), “told only false” (**f**), “told both true and false” (**b**), and “told neither true nor false” (**n**). These values, ordered by the  $\leq$  relation, form a “logical lattice” (figure 6.2).

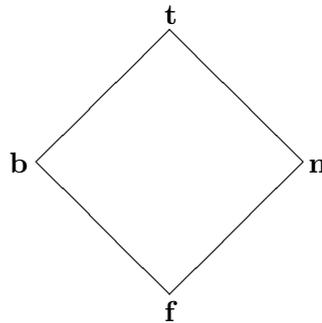
Are there any permutations of  $V$  that preserve this structure? Yes, there are two: the

---

<sup>38</sup>See Dummett 1991:43–4. Typically, the multivalues form a *lattice* with respect to g.l.b. and l.u.b. on the relation  $\leq$ , but we need not require even that  $\leq$  be a partial ordering. For example, antisymmetry will be violated when there are several designated multivalues, each of them related by  $\leq$  to each of the others. If  $\leq$  is not transitive, logical consequence will not be transitive either, but I see no reason to build in any such requirement.

<sup>39</sup>When the multivalues form a lattice, logical consequence can be defined as follows: an argument from a set  $\Gamma$  of sentences to a sentence  $S$  is *valid* just in case on every interpretation, the meet (g.l.b.) of the semantic values assigned to the sentences in  $\Gamma$   $\leq$  the semantic value assigned to  $S$ .

Figure 6.2: The logical lattice for a four-valued logic.



trivial identity permutation and the permutation that switches **b** and **n** without changing **t** or **f**. It turns out that the matrices Belnap and Anderson give for *conjunction*, *disjunction*, and *negation* are all invariant with respect to these permutations (no surprise, since these are just the meet, join, and complement operations on the logical lattice).<sup>40</sup>

Third, consider the nine-valued logic whose values represent the reports of two sources (A and B) that a proposition is true, false, or unknown (not yet reported).<sup>41</sup> For example, the value **nt** represents a state of information in which B has reported the proposition to be true, while A has not yet made any report. As before, these values form a lattice under

<sup>40</sup>In fact, out of the  $2^{32}$  possible binary four-valued matrices, only  $2^{16}$  are invariant under a transformation that switches **b** with **n** and maps **t** and **f** to themselves. Proof: Number the squares of the matrices as follows:

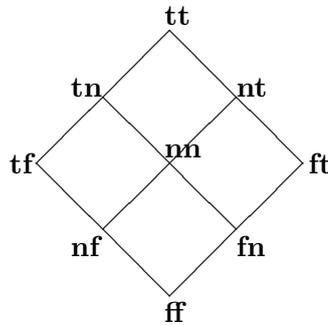
	<b>n</b>	<b>f</b>	<b>t</b>	<b>b</b>
<b>n</b>	1	2	3	4
<b>f</b>	5	6	7	8
<b>t</b>	9	10	11	12
<b>b</b>	13	14	15	16

It is easy to see that squares 6, 7, 10, and 11 must contain **ts** or **fs**, if the matrix is to be invariant. The values in the other squares can be changed in independent pairs: 1 and 16, 4 and 13, 5 and 8, 9 and 12, 2 and 14, 3 and 15. (In these pairs,  $\langle x, y \rangle$  is linked with  $\langle \sigma^{-1}(x), \sigma^{-1}(y) \rangle$ .) Each pair must contain either  $\langle \mathbf{t}, \mathbf{t} \rangle$ ,  $\langle \mathbf{f}, \mathbf{f} \rangle$ ,  $\langle \mathbf{n}, \mathbf{b} \rangle$ , or  $\langle \mathbf{b}, \mathbf{n} \rangle$ , if the matrix is to be invariant. So there are  $2^4$  possible inner  $2 \times 2$  squares, each with  $4^6$  possible outer shells (6 independent pairs with 4 possible values each), for a total of  $(2^4)(4^6) = 2^{16}$  invariant matrices in all.

<sup>41</sup>For some purposes it might be useful to add the possibility of contradictory reports (both true and false), but the simpler logic considered here suffices to make my point. Anderson and Belnap 1992 consider the possibility of multivalued values that track the source of reports, but they do not pursue it (523).

the implication-relevant  $\leq$  relation (figure 6.3).

Figure 6.3: The logical lattice for a nine-valued logic.



There is one non-trivial permutation of  $V$  that preserves this structure: the symmetric flip across the vertical axis, which flips **tn** with **nt**, **tf** with **ft**, and **nf** with **fn**, mapping **tt**, **nn**, and **ff** to themselves. (If we consider the interpretation of the values, this transformation amounts to switching the reports of A and B.) As before, the logical operators (conjunction, disjunction, and negation, construed as the meet, join, and complementation operations on the lattice) are invariant with respect to this permutation.<sup>42</sup> The noninvariant matrices are those that treat A's and B's reports differentially; it seems reasonable to require that logical operators not do this.

By taking the implicational ordering  $\leq$  as intrinsic structure on a type  $V$  of multivalues, then, we obtain a nontrivial criterion for logicality in the types derived from  $V$ . The criterion is intuitively plausible, both because it certifies as logical the standard logical operators in multivalued logics and because the structure it invokes—the implicational ordering on multivalues—is central to the aims of logic. So the structure imposed on  $V$  by  $\leq$  is a plausible *prima facie* candidate for intrinsic structure. What we still lack is a *general* characterization of intrinsic structure. Ideally, we would like to be able to say what it is for structure on a presemantic type to be intrinsic, in a way that explains why the structure imposed by  $\leq$  is intrinsic on  $V$  (assuming that it is) and why it is appropriate for *logical* notions to be

<sup>42</sup>Of the  $3^{162}$  possible nine-value matrices for binary connectives,  $3^{81}$  of them are invariant. The numbers are calculated as in note 40.

sensitive to intrinsic structure. Until we can do this, we have no grip on the connection between logicity and invariance under all intrinsic structure-preserving permutations of the basic types, and hence no understanding of why the invariance criterion should count as a criterion for *logicality*. In the interest of greater generality, then, let us consider how the invariance criterion might be extended to apply to intensional operators.

## 6.5 Extending the account to intensional logics

Sher explicitly excludes modal logic from the scope of her investigation:

First, my criterion for logical terms is based on an analysis of the Tarskian framework, which is insufficient for modals. Second, we cannot take for granted that the task of modal logic is the same as that of symbolic logic proper [*sic*]. To determine the scope of modal logic and characterize its operators, we would have to set upon an independent inquiry into its underlying goals and principles. (1991:54)

The conception of logic Sher articulates is “not intended to detract from the value of intensional logics” (1996:679 n. 30). But it does, on her view, “point to a difference between the philosophical principles underlying mathematical logic and those underlying the various intensional logics” (ibid.). This limitation of the scope of the invariance criterion significantly reduces its interest. Even if we are interested primarily in extensional languages, we want to know what makes terms in these languages *logical*, not what makes them *extensional-logical*. Moreover, in view of the strong formal analogies between quantifiers and modal operators, it would be surprising if criteria for the logicity of extensional quantifiers could not be fruitfully extended to modal operators.

Granted, such an extension would be difficult to achieve in Sher’s Tarskian semantic framework, which would require rather extensive modification in order to make room for intensional operators. But the extension is obvious and intuitive in the functional-categorical framework we have been using throughout this chapter. In fact, there are two ways in which it might be accomplished.

The first is to take  $V$  to be a type of “propositions” (conceived as functions from possible worlds to truth values, or more simply as sets of possible worlds) and take values for the modal and other sentential operators from  $(V \Rightarrow V)$ ,  $(V \Rightarrow (V \Rightarrow V))$ , and so on. All that is then needed in order to extend the invariance criterion (as reformulated in the section 6.4) to modal logic is a specification of the intrinsic structure on  $V$ . The modal operators will be those items in  $(V \Rightarrow V)$ ,  $(V \Rightarrow (V \Rightarrow V))$ , etc., that are invariant under all permutations of  $V$  that preserve intrinsic structure. (As before, a reasonable candidate for semantic structure would be the  $\leq$  relation on  $V$ .)

On the second approach (van Benthem 1989:334), which emphasizes the analogies between modal operators and quantifiers rather than the fact that modal operators are sentential operators, we keep the type  $V$  two-valued and add a *third* basic type  $W$  (for “worlds”).<sup>43</sup> Semantic values for sentences can then be drawn from  $(W \Rightarrow V)$ , values for unary sentential connectives from  $((W \Rightarrow V) \Rightarrow (W \Rightarrow V))$ , values for one-place predicates from  $(O \Rightarrow (W \Rightarrow V))$ , values for first-order quantifiers from  $((O \Rightarrow (W \Rightarrow V)) \Rightarrow (W \Rightarrow V))$ , and so on.<sup>44</sup> The invariance criterion applies immediately to this framework: the logical notions in a type  $Z$  are just those that are invariant under all permutations of the basic types from which  $Z$  is derived.

Here we will pursue the second approach. Which unary modal operators—which items in  $((W \Rightarrow V) \Rightarrow (W \Rightarrow V))$ —are permutation invariant? That depends on what we count as intrinsic structure on  $W$ . If we take the intrinsic structure on  $W$  to be the null structure, then invariant unary modal operators include

- the identity operator
- truth-functional negation: the operator that takes a function  $f$  from possible worlds

---

<sup>43</sup>Here I am assuming, for the sake of simplicity, that the set of worlds is given once and for all and does not vary from one interpretation to another. Standard semantic practice is to take the set of worlds as given by a *frame* or *model structure* that may differ from one interpretation to another. This complication can be handled in the same way as the use of variable quantifier domains (see section 6.3, above): that is,  $W$  can be taken to be an *indexical* type that determines different sets of worlds and accessibility relations in different contexts.

<sup>44</sup>Other choices are possible here: for instance, we might take use items in  $(W \Rightarrow O)$  rather than  $O$  as semantic values for singular terms (individual concepts), and similarly items in  $((W \Rightarrow O) \Rightarrow (W \Rightarrow T))$  for predicates (Bressan-style intensional predication).

to truth values to another such function  $f^*$ , such that for all  $w \in W$ ,  $f^*(w) = \text{True}$  iff  $f(w) = \text{False}$ .

- a necessity operator, which takes *Necessarily-True* (the function that takes the value True on every possible world) to *Necessarily-True* and all other functions in  $(W \Rightarrow V)$  to *Necessarily-False* (the function that takes the value False on every possible world).
- a possibility operator, which takes *Necessarily-False* to *Necessarily-False* and all other functions in  $(W \Rightarrow V)$  to *Necessarily-True*.
- the negations of the necessity and possibility operators
- an infinite number of modal analogues to the cardinality quantifiers: e.g., the operator that takes every function in  $(W \Rightarrow V)$  which takes the value True on exactly five possible worlds to *Necessarily-True* and every other function in  $(W \Rightarrow V)$  to *Necessarily-False*.

This yield of logical unary operators for modal logic both exceeds and falls short of the usual selection. It goes beyond the usual modal logics in its inclusion of analogues to the cardinality quantifiers. But this affords no grounds for objection to the criterion as a *necessary* condition for logicality. Moreover, although it is difficult to imagine a use for the modal cardinality operators, it does not seem intuitively wrong to count them as logical in a language that has them. Indeed, similar logical constants can be found in metric tense logics: e.g., “five hours ago, it was the case that...”

On the other hand, the criterion’s yield of logical notions is in certain respects *more* restrictive than the usual logical practice admits. As formulated above, the criterion leaves no room for logical operators defined—as modal operators usually are—in terms of an *accessibility relation* on the set of possible worlds. For the accessibility relation will not (in general) be preserved by all permutations of  $W$ . For example, let  $L$  be the function in  $((W \Rightarrow V) \Rightarrow (W \Rightarrow V))$  such that for all  $w \in W$  and  $f \in (W \Rightarrow V)$ ,  $L(f)(w) = \text{True}$  just in case  $f(w')$  is true for every  $w'$  accessible from  $w$ . ( $L$  is of course the necessity operator as it is usually defined in modal logic.) Suppose that  $W = \{w_1, w_2\}$ , where  $w_1$  is the only world

accessible from  $w_2$  and both  $w_1$  and  $w_2$  are accessible from  $w_1$ . Let  $\sigma$  be a permutation of  $W$  that switches  $w_1$  with  $w_2$ . Then  $L$  is not invariant under  $\sigma$ . To see why, let  $f(w_1) = \text{False}$  and  $f(w_2) = \text{True}$ . Then  $L(f)(w_1) = \text{False}$ , but  $\sigma^{((W \Rightarrow V) \Rightarrow (W \Rightarrow V))}(L)(f)(w_1) = \text{True}$ .<sup>45</sup> So  $L$  is not invariant under  $\sigma$ .

There is one special case in which  $L$  will be invariant under all permutations of  $W$ : the case where the accessibility relation on  $W$  is *universal*, that is, where any possible world is accessible from any other. (Here  $L$  will be the necessity operator ruled logical in the discussion of the invariance criterion above.) Since validity for the modal logic S5 can be defined as validity on every frame in which the accessibility relation is universal, we might be tempted to conclude that only the S5 operators are really logical, according to the invariance criterion.<sup>46</sup> The non-S5 modal operators (those of S4, B, K, and so on) are sensitive to structure on  $W$  that is not preserved by arbitrary permutations. Similarly, (linear) tense operators like “it was the case that...” are not invariant with respect to permutations of the type  $T$  of *times* that fail to preserve the structure of temporal succession.

It might be argued that this narrowing of the scope of logic is demanded by the fundamental intuitions behind the invariance criterion. As we have seen, the invariance criterion rules as non-logical any extensional operator that respects relations holding between particular objects. The  $\in$  of set membership, for instance, is outside the bounds of logic because it respects the set-theoretic membership relation that holds between, e.g., 1 and  $\{1, 2\}$ . So it might seem reasonable to demand on analogous grounds that logical *intensional* operators not respect relations holding between possible worlds or times, including relations of acces-

---

<sup>45</sup>Proof:  $\sigma^{((W \Rightarrow V) \Rightarrow (W \Rightarrow V))}(L)(f) = (\sigma^{(W \Rightarrow V)} \circ L \circ (\sigma^{(W \Rightarrow V)})^{-1})(f) = (\sigma^{(W \Rightarrow V)} \circ L \circ (\sigma^{-1})^{(W \Rightarrow V)})(f)$  [by the inverse theorem, section 6.1.4, above]  $= (\sigma^{(W \Rightarrow V)} \circ L)((\sigma^{-1})^{(W \Rightarrow V)}(f)) = (\sigma^{(W \Rightarrow V)} \circ L)(f \circ (\sigma^{-1})^{-1}) = (\sigma^{(W \Rightarrow V)} \circ L)(f \circ \sigma) = \sigma^{(W \Rightarrow V)}(L(f \circ \sigma)) = L(f \circ \sigma) \circ \sigma^{-1}$ . Now  $(f \circ \sigma)(w_1) = f(w_2) = \text{True}$  and  $(f \circ \sigma)(w_2) = f(w_1) = \text{False}$ . So  $L(f \circ \sigma)(w_1) = \text{False}$  and  $L(f \circ \sigma)(w_2) = \text{True}$ . So  $\sigma^{((W \Rightarrow V) \Rightarrow (W \Rightarrow V))}(L)(f)(w_1) = (L(f \circ \sigma) \circ \sigma^{-1})(w_1) = L(f \circ \sigma)(w_2) = \text{True}$ . (Here I am assuming that  $\sigma^V$  is the identity permutation.)

<sup>46</sup>Usually S5 validity is defined as validity in every frame in which the accessibility relation is transitive, symmetric, and reflexive. But this is equivalent to validity in every frame with a universal accessibility relation (Hughes and Cresswell 1968:61). It should be noted, however, that the difference between frames with universal accessibility relations and frames with transitive, symmetric, and reflexive, but not universal accessibility relations is significant for assessing invariance. For the modal operators will not be invariant unless the accessibility relation is universal.

sibility or temporal succession. And that means that the constants of S4 and tense logic are not logical. Some proponents of the invariance criterion have swallowed this consequence (McCarthy 1981:513, Scott 1970:161).

But we need not swallow it—at least, not yet. The discussion in section 6.4 shows us how to resist the demand that logical intensional operators be insensitive to *all* particularity in a type (in this case, the set  $W$  of possible worlds). For no one would say that logical operators must be insensitive to individual characteristics of *truth values* (the items in  $V$ ). We can demand that logical notions be invariant under permutations of the basic type  $V$  only if we restrict the allowable permutations to those that preserve some *intrinsic structure* on  $V$ . If the type  $O$  of objects has the null intrinsic structure, then that is a peculiar feature of  $O$ , not one we can assume to be shared by  $W$ . Indeed, the fact that non-S5 modalities are widely accepted as logical suggests that the accessibility relation on  $W$  ought to be regarded as intrinsic structure, just like the  $\leq$  relation on multivalued. If we regard the accessibility relation as intrinsic structure, then a necessity operator defined in the usual way ( $L\phi$  is true at  $w$  iff  $\phi$  is true at every world accessible from  $w$ ) will count as logical, even if we allow frames with non-universal accessibility relations.

For  $W$  as for  $V$ , then, we have a plausible candidate for intrinsic structure, but we still lack a general theoretical characterization of intrinsic structure, by reference to which we might justify our choices. Why, for instance, should the accessibility relation count as intrinsic structure on  $W$ , but not the relation that holds between two worlds just in case the first contains more sheep than the second? This is just another variation of the *justificatory challenge* that we faced in the previous section (page 209, above). Without some general characterization of intrinsic structure, our choices are to a certain extent *ad hoc*: they are motivated by our antecedent intuitions that certain notions are logical.<sup>47</sup> This would be unobjectionable if our only concern were to *systematize* these antecedent intuitions, but proponents of the invariance criterion generally have something more ambitious in

---

<sup>47</sup>Van Benthem 1989:334, who acknowledges the possibility of demanding that permutations of  $W$  preserve a certain structure, notes that “... the systematic question then becomes how to motivate (a minimum of) such additional structure independently.”

mind. The aim is to explain why certain notions are logical on the basis of theoretical considerations about logicality, and this aim is undermined unless we can connect the notion of intrinsic structure to these theoretical considerations.

Finally, it should not be thought that the difficulty of giving a non-*ad hoc* justification for privileging a certain structure on  $W$  supports the view that *no* structure on  $W$  should be privileged as “intrinsic.” From the present perspective, the null structure is just one choice among many possible choices for intrinsic structure. It stands in need of justification as much as any other choice.

## 6.6 Limitations of the invariance criterion

In the last two sections, we have generalized the invariance criterion by considering transformations of *all* the basic presemantic types, not just the type  $O$  of objects. We have seen how the criterion can be usefully applied to the semantic values found in sentential and intensional logics, as well as to extensional quantifiers. What we have learned from generalizing the account in this way is that we must regard *all* the semantic types (including  $O$ ) as having an “intrinsic structure.” And now we are in a position to see just how little philosophical bread the invariance criterion bakes. For, granting that the invariance criterion provides a plausible answer to the question:

Given a set of basic presemantic types (including their intrinsic structures),  
which items in the types derived from these are logical?

the answer to this question won't be much help unless we can answer a second one:

In virtue of what does some structure on a type count as *intrinsic*?

But proponents of the invariance criterion say nothing to answer this question. Rather than *arguing* that the intrinsic structure of  $O$  is the null structure, they simply *presuppose* this. And many of the positive conclusions about logical constants that they draw from the invariance criterion depend as much on this unargued presupposition as on the appeal to invariance.

The case of  $\in$  (the set-theoretic membership relation) illustrates this point well. There are serious questions about the extent to which set theory is logical, and both Tarski (1966:153) and Sher (1991:58) take the invariance criterion to address them. On their view, the invariance criterion rules  $\in$  non-logical. It does so, however, only under the assumption that the intrinsic structure of  $O$  is the null structure. And unless this assumption can be further justified, the argument against the logicity of  $\in$  simply begs the question, since the invariance criterion would certify set-theoretic membership as a logical notion if we took set membership relations on the type  $O$  as intrinsic structure.

A similar point can be made about the mereological sum operator, which takes a sequence of individuals to their mereological sum.<sup>48</sup> Whether this operator is permutation-invariant depends on whether the mereological part/whole relations on  $O$  count as intrinsic structure. Apart from some assumption about the intrinsic structure of  $O$ , then, the invariance criterion says nothing about the logicity of the sum operator.

As van Benthem 1989 points out, even the quantifier “all blondes” is invariant under permutations that respect bloneness (320). Indeed, if we held that relations of color complementation were logical, we could add this structure, too, to  $O$ , excluding permutations that switched red things and blue things without also switching green things and orange things. By itself, then, the invariance criterion does not even rule out a “logic of color terms.”

The upshot is that unless we can give an independent characterization of intrinsic structure, the invariance criterion gives very little guidance about which notions are logical. For we can always broaden the class of semantic values that count as logical by countenancing more intrinsic structure on the basic categories. At most the invariance criterion will rule out certain *combinations* of verdicts on logicity, preventing us from holding, for instance, that set-theoretic membership is a logical notion in some system in which the subset relation is non-logical.

---

<sup>48</sup>This operator is sometimes taken to be a logical constant: see Massey 1976 for an argument.

Have any of the proponents of the invariance criterion said anything that might be used to fill this lacuna in their account—that is, to justify their assumption that the intrinsic structure of the type  $O$  is the null structure? Let's look at how they motivate their accounts.

As we have seen (page 175, above), Tarski 1966 presents his invariance criterion in the context of Klein's Erlangen program. The permutations on a domain constitute the *maximal* class of one-one mappings from  $O$  into itself. But if maximality *alone* were a reason for demanding invariance under *all* permutations on  $O$ , it ought equally to be a reason for demanding invariance under all permutations on  $V$ . And not even conjunction is invariant under *all* permutations of a two-valued  $V$ . Hence the motivation for taking  $O$  to have the null intrinsic structure must go beyond mere maximality, but here Tarski is silent.

Sher 1991 (43) elicits the invariance criterion from the notion of formality to which Tarski appeals in his 1936 article on logical consequence:

...since we are concerned here with the concept of logical, i.e. *formal*, consequence, and thus with a relation which is to be uniquely determined by the form of the sentences between which it holds, this relation cannot be influenced in any way by empirical knowledge, and in particular by knowledge of the objects to which the sentence  $X$  or the sentences of the class  $K$  refer. (Tarski 1936:414-15)

Perhaps it would violate the spirit of this condition to allow structure of an *empirical* kind—perhaps even color complementation relations—to count as intrinsic on the type  $O$ . But Tarski would not have considered set-theoretic inclusion and mereological parthood to be *empirical* relations.

McCarthy 1981 motivates the invariance criterion as a precise way of capturing the traditional desideratum that logic be *topic-neutral* (508-9). Perhaps it is clear that color complementation relations are not topic-neutral, and on that ground we might disqualify them as intrinsic structure. But in what sense do relations of set-theoretic inclusion fail to be topic-neutral? *Anything* can be (in fact, is) a member of sets. To say that set theory is not topic-neutral on the grounds that it concerns a particular kind of object—sets—will seem right only if we are antecedently committed to a construal of topic-neutrality as permutation invariance. Set-theoretic notions are certainly applicable to discourse about any topic: why

isn't that enough to qualify them as topic-neutral?<sup>49</sup> Thus McCarthy really hasn't given any justification for his (implicit) decision not to count set-theoretic membership relations on  $O$  as intrinsic structure.

I am not arguing that set-theoretic membership ought to count as a logical notion. I am simply pointing out that proponents of the invariance criterion have said nothing to justify their assumption that the type  $O$  has the null intrinsic structure. Yet it is on this assumption that their most significant negative results—the denial that  $\in$  is a logical constant, the denial that there are any logical singular terms, the denial that the S4 necessity or past tense operators are logical—largely depend. It is this assumption that is baking the bread, not the appeal to invariance.

This criticism can be cast in two different forms, depending on how we think about the basic presemantic types. If we think of the types as specifiable independently of their intrinsic structure (so that we can talk of  $O$ , for instance, without assuming that its intrinsic structure is the null structure), then the criticism is that we have only an *immanent* definition of intrinsic structure (in the sense of Evans 1976), where a *transcendent* one is needed. That is, we know what it is for the invariance criterion to *treat* some structure on a presemantic type as intrinsic, but we have not yet said anything about which structure or structures it is *correct* to treat as intrinsic. We can weigh various choices of intrinsic structure against each other by comparing the output of the invariance criterion (given those choices) to our intuitions about particular cases, but we had hoped for a kind of *explanation* of logical constanhood, not just a systematization of our intuitions. What is needed to satisfy these hopes is a transcendent definition of intrinsic structure, which would tell us which structures on the basic presemantic types are *really* intrinsic.

If, on the other hand, we think of the basic presemantic types as *individuated* in part by their intrinsic structures, then the question of what structure on a type counts as intrinsic becomes trivial: for one has not picked out a determinate type about which to ask the

---

<sup>49</sup>Similar points hold for the basic notions of mereology and even of arithmetic. See section 3.5, above.

question until one has specified its intrinsic structure. On this view, a type is not just a set of elements, but a set of elements plus an intrinsic structure. Strictly speaking, then, there is not just one type  $O$  of objects, but many  $(O_1, O_2, O_3, \dots)$ , which differ only in their intrinsic structure. In this case, it can no longer be charged that the notion of intrinsic structure lacks a transcendent definition: given a basic type, it is perfectly definite what counts as its intrinsic structure, and the invariance criterion yields a perfectly definite criterion for the logicity of notions in the types derived from it. The problem now is not which structure on  $O$  is intrinsic, but rather which of the many structured basic types  $O_i$  we should use in the semantics for a quantificational language. The set-theoretic membership relation will be invariant on  $O_1$ , but not on  $O_2$ ; the mereological sum function will be invariant on  $O_3$ , but not on  $O_4$ ; and so on. In effect, we have avoided the problem of giving a transcendent definition of intrinsic structure only by making the invariance criterion *relational*. The invariance criterion will not supply an answer the question, “is  $\in$  a logical constant?” until we have chosen one of the  $O_i$ 's. And on what basis should we make the choice? Is it ever appropriate to use a type  $O$  with non-trivial intrinsic structure?

It does not much matter in which of these ways we pose the problem. In either case, the moral is the same: the invariance approach's invocation of 2-formality does not by itself generate a definite criterion for logical constants. Something more is needed: either a transcendent definition of intrinsic structure or a principle for selecting the appropriate structured presemantic type to use in the semantics for a language. In the next section, I will argue that we can bridge this gap by appealing to 1-formality.

## 6.7 Intrinsic structure: a proposal

For the sake of definiteness, I will pose the problem in the second of the two ways discussed above: what makes it appropriate, in formalizing the semantics for a language, to use a presemantic type with a particular intrinsic structure? I approach this problem by distinguishing between two sources of constraints on presemantic types: semantics proper

and postsemantics. Constraints on intrinsic structure, I will suggest, should come from postsemantics:

**Intrinsic Structure Principle.** One should use a type with just as much intrinsic structure as is needed for the postsemantics, and no more.

Notions invariant under permutations that preserve this structure have a special connection to the most general features of contents: their capacities to be used in assertion or inference. Hence the laws governing these notions can reasonably be claimed to be normative not just for particular kinds of conceptual activity (for instance, thought about the physical world), but for conceptual activity as such. I will show how this suggestion can be applied in several “test cases”: multivalued sentential logics, modal logics, tense logics, and extensional logic.

### 6.7.1 Two sources of constraints on presemantic types

The subdiscipline I have dubbed “presemantics” concerns itself exclusively with semantic values and their types, abstracting from the linguistic expressions to which these semantic values will eventually be assigned. Nonetheless, presemantics gets its point from the larger endeavor of which it is a part: the construction of a formal semantics for a language. This larger endeavor determines which presemantic types ought to be articulated in the presemantics. Constraints on types come from two sources: semantics proper and postsemantics.<sup>50</sup>

The project of semantics proper is to show how the semantic values of grammatically complex expressions depend on the semantic values of their parts. Which semantic values—and hence which basic types—are needed for this purpose depends on the expressive capacities of the language. For example, if the only sentential operators the language is to have are the standard truth-functional ones, then a two-valued type suffices to supply semantic values for sentences. On the other hand, if the language is to contain sentential operators that track epistemic features of sentences (for example, “told true only,” or “told

<sup>50</sup>For this taxonomy of subdisciplines, see section 6.2.1, above.

true by A and false by B”), then we will need to draw semantic values for sentences from a multivalued type (as in the four- and nine-valued logics discussed in section 6.4, above). Similarly, if the language is to have modal sentential operators (whether or not they are counted as logical), then a finite-valued type will not supply enough semantic values for a compositional semantics. Instead, we will need a type containing infinitely many *propositions*. (Alternatively, we can keep V two-valued, add a type W of possible worlds, and assign sentences semantic values from the derived type  $(W \Rightarrow V)$ .) The more expressive power our language is to have, the finer we will have to chop our semantic values in order to give a compositional semantics. It is worth emphasizing that at this point, no distinction need be made between logical and non-logical expressive power. One might want to draw sentential semantic values from a type containing propositions even in a language whose only *logical* constants are truth-functional, in order to accommodate *non-logical* intensional operators.

Postsemantics has a different goal. It assumes that the project of semantics has been accomplished, so that given an assignment of semantic values to the language’s grammatically simple terms, the semantic values of all of its complex terms are determined as well. Postsemantics concerns the *import* of these values: that is, their relation to the fundamental or “top-level” semantic notions—generally truth and implication—that must be invoked in explaining the *use* of language.<sup>51</sup>

The distinction between semantics proper and postsemantics is somewhat obscured in the case of classical two-valued logic, because the same classification of sentences (into true and false) that is used in giving a compositional semantics connects directly to the use of language—to proprieties for correct assertion. But as Dummett has emphasized (1959, 1973:ch. 12), there are really two distinct notions of truth value. The first notion applies to “sentences as complete utterances” (1973:417) and is directly relevant to an explanation

---

<sup>51</sup>There is a tradition, going back at least to Frege, of taking truth and implication, the semantic notions most directly connected with proprieties for assertion and inference, as fundamental for the explanation of *all* use of language. It may be that this tradition is misguided, and that other top-level notions are required in order to understand the proprieties for non-assertional uses of language, e.g., questions. If so, then they, too, will belong to postsemantics. My talk of truth and implication as the fundamental notions of post-semantics should be understood as illustrative, not exclusive. (For a polemic against the focus on declaratives as a paradigm of language use, see Belnap 1990).

of what is accomplished by the assertion of the sentences. (One who sincerely utters an assertoric sentence intends to speak the truth, can be held responsible for factual error in the event that the sentence is false, and so on.) The second notion of truth value applies to sentences as (potential) components of more complex sentences. This “ingredient” truth value is relevant only to the determination of the “stand-alone” truth value of a complex sentence; it is only indirectly connected to proprieties of use (via its contribution to the stand-alone truth value). In our terms, the first notion of truth value belongs to postsemantics, while the second belongs to semantics proper. As Dummett notes, “[t]here is no a priori reason why the two notions of truth-value should coincide,” and in multivalued logics, they do not. The multivalued values are necessary in order to give a compositional semantics, but when it comes to explaining the proprieties for the use of sentences as complete utterances, only the distinction between designated and undesignated values matters:

Thus an assertion made by uttering a given sentence amounts to a claim that that sentence has a designated value: in order, therefore, to grasp the content of any particular assertion, all that is necessary is to know the condition for the sentence uttered to have a designated value. We do not need, for this purpose, to know anything about the distinction between the different designated truth-values or between the different undesignated ones: an understanding of those distinctions is required only in order to be able to derive the assertibility condition of a complex sentence from the senses of its constituents, since the semantic roles of the sentential operators are given by truth-tables which relate to the individual truth values and not just to the distinction between a designated and an undesignated value. (Dummett 1973:422-3)<sup>52</sup>

A similar distinction can be made in modal logic. If there is to be a compositional semantics for a modal language, then the semantic values of sentences must be individuated much more finely than the truth values. But when it comes to assessing the truth or falsity of a stand-alone utterance, all that matters is its truth value at the world of utterance.

The distinction between postsemantics and semantics proper is more prominent, even in

---

<sup>52</sup>I use Dummett as an example of how one might draw a distinction between semantic and postsemantic levels; I do not mean to endorse Dummett’s view as a *general* claim about multivalued logics. Surely one might (for certain purposes) invoke multivalued values in explaining proprieties for the use of sentences. My point is only that distinctions between semantic values *need* not have any role at the postsemantic level. (Thanks to Nuel Belnap here.)

classical logic, if we attend to proprieties of *inference* rather than assertion. It is clear that when we have given an account of how the semantic values of complex expressions depend on the semantic values of their parts, we do not yet have an account of implication. Something more is required: the definition of implication as truth-preservation in all interpretations (in classical languages), or as the preservation of some relation ( $\leq$ ) defined on the multivalues (in multivalued languages), or as the preservation of truth at every context of utterance on every interpretation (in modal languages). This further step belongs to postsemantics.<sup>53</sup>

The role of postsemantics, then, is to mediate between the semantic values required for the purposes of compositional semantics and the fundamental semantic notions in terms of which the use of language (e.g., proprieties of assertion and inference) is to be explained. The task of postsemantics imposes further constraints on the basic semantic types, beyond the constraints imposed by semantics proper. For example, in a multivalued logic, there must be a distinction between designated and undesignated values in  $V$  (or some other way of going from ingredient truth values to stand-alone truth values). And there must be a relation  $\leq$  on multivalues by means of which implication can be defined. (Sometimes the distinction between designated and undesignated can perform this function, but not always.)

### 6.7.2 What is intrinsic structure?

My suggestion (the Intrinsic Structure Principle) is that the intrinsic structure on the basic presemantic types ought to be just the structure required for the purposes of postsemantics: for example, the distinction between designated and undesignated multivalues, or the relation  $\leq$  on multivalues. That is, in picking presemantic types for a particular semantic purpose, we should choose types with just as much intrinsic structure as is required by the postsemantics (and no more). I will consider several examples below, but for now,

<sup>53</sup>Note that in *supervaluational* semantics, a precisely analogous further step will be required in order to give an account of the *truth* of stand-alone assertions: a sentence is true (in the sense directly relevant to proprieties of assertion) just in case it receives the semantic value True in every interpretation in a certain class.

consider a language that requires only a single two-valued type ( $V$ ) for the purposes of semantics proper (that is, a language in which the only basic terms are atomic sentences and truth-functional sentential operators). If we are to connect these values with proprieties for assertion, we must mark one (True) as designated and the other (False) as undesignated; and if we are to connect them with proprieties for inference, we must recognize an ordering on them ( $\text{False} \leq \text{True}$ ) by means of which implication can be defined. So we need a type  $V$  in which both these structures—the distinction between designated and undesignated and the relation  $\leq$ —are intrinsic. (This type will admit of no nontrivial intrinsic structure-preserving permutations, and so every function on  $(V \Rightarrow V)$ ,  $(V \Rightarrow (V \Rightarrow V))$ , etc.—that is, every truth function—will be invariant.)

By conceiving intrinsic structure in this way, we acquire an intelligible rationale for demanding that logical notions be invariant only under permutations that preserve intrinsic structure. Sensitivity to *intrinsic* structure does not compromise the general applicability of a logical notion, because intrinsic structure belongs to a type in virtue of the most general purpose of logical theory: the study of the semantic relationships that hold between sentences solely in virtue of their capacity for being asserted and used in inferences. On this ground, one might say that notions that are sensitive only to intrinsic structure are applicable to *thought as such*, independent of its particular subject matter, and that the laws governing these notions are norms for thought as such.

Thus the invariance criterion, combined with the account of intrinsic structure offered above, yields a conception of logic as *1-formal*. This is surprising, because we began by thinking of the invariance criterion as a way of spelling out the notion of *2-formality* apart from its more obscure cousins. What has emerged is that the mere appeal to invariance does not yield a determinate conception of logic at all until something is said about intrinsic structure. And in order to say something about intrinsic structure, we must bring to bear considerations that go beyond mere permutation invariance. In understanding intrinsic structure as the structure that must be recognized on a presemantic type in an account of the top-level semantic notions (e.g., truth and implication)—notions that are applicable

to sentences solely in virtue of their being potential vehicles of assertion and inference (or more generally, of language use), and independently of their particular content—I have drawn from the tradition of conceiving logic as 1-formal, that is, as normative for thought or conceptual activity as such.

Let us now consider how the notion of intrinsic structure as structure needed for the purposes of postsemantics plays out in some of the test cases we have already considered.

### 6.7.3 Application: multivalued logics

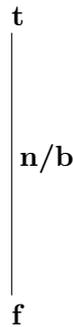
Consider first the four-valued logic discussed in section 6.4, above. Our account of implication (in the postsemantics) must make reference to the  $\leq$  relation defining the “logical lattice” (figure 6.2, page 212, above). Let us take this relation as intrinsic structure on  $V$ . As we saw earlier, this choice of intrinsic structure allows just one nontrivial intrinsic structure-preserving permutation: the one that permutes  $\mathbf{n}$  and  $\mathbf{b}$ . Given this choice of intrinsic structure, the invariance criterion says, in effect, that logical notions must not treat  $\mathbf{n}$  differently from  $\mathbf{b}$ . We can now say why this demand is appropriate. Unlike the difference between  $\mathbf{t}$  and  $\mathbf{f}$ , or between  $\mathbf{t}$  or  $\mathbf{f}$  and  $\mathbf{n}$  or  $\mathbf{b}$ , the difference between  $\mathbf{n}$  and  $\mathbf{b}$  is irrelevant for the assessment of inferential proprieties between stand-alone sentences. It plays no role in the postsemantic account of implication.

This is not to say that  $\mathbf{n}$  and  $\mathbf{b}$  should be identified, resulting in a three-valued type  $V$  with no non-trivial structure-preserving permutations (see figure 6.4).<sup>54</sup> The two values play different roles in the compositional semantics, and they may be treated differently by sentential operators. Moreover, the fact that there are *two* incomparable values between  $\mathbf{t}$  and  $\mathbf{f}$  is important even in the postsemantics. An interpretation that assigns  $\mathbf{b}$  to a sentence  $A$  and  $\mathbf{n}$  to a sentence  $B$  witnesses the failure of implication from  $A$  to  $B$ , while an interpretation that assigned the combined value  $\mathbf{n/b}$  to both  $A$  and  $B$  would not. But although the postsemantics is sensitive to the fact that there are two incomparable values between  $\mathbf{t}$  and  $\mathbf{f}$ , it does not care which is which. Whereas the difference between  $\mathbf{t}$  and

<sup>54</sup>See Anderson and Belnap 1992:522-3 on the temptation to identify  $\mathbf{n}$  and  $\mathbf{b}$ .

**b** is semantically significant no matter which sentential operators are present in the language, the difference between **n** and **b** gets its semantic significance only indirectly (if at all), through the presence of sentential operators that are sensitive to it. If the language contained no operators sensitive to the difference between **n** and **b**, this difference would be semantically irrelevant, and the two values could be (uniformly) switched without affecting the postsemantics. The difference between **t** and **b**, on the other hand, would be semantically relevant even if the language contained only atomic sentences. Thus the invariance criterion (together with the Intrinsic Structure Principle) captures the intuition that logical notions should be sensitive only to differences whose semantic significance does not depend on the particular expressive power of a language (that is, on which expressions it contains in each grammatical category).

Figure 6.4: Four values collapsed into three.



This point is even more evident in the case of the nine-valued logic considered in section 6.4. At the postsemantic level, in determining which stand-alone sentences imply which, the difference between A's and B's reports is irrelevant (although the fact that there are two reporters is not). If we switched A's reports with B's, all of the same implications between stand-alone sentences would hold. The difference between A's and B's reports is relevant only in the compositional semantics, and its relevance is contingent on the presence of operators that are sensitive to this difference.<sup>55</sup> In requiring that logical notions be

<sup>55</sup>For example, a unary operator that negates A's report without changing B's, taking **tf** to **ff**, **fn** to **tn**, **nn** to **nn**, and so on.

insensitive to this difference, the invariance criterion captures the idea that logical notions abstract from the particular semantic content of a language.

#### 6.7.4 Application: modal logics

In our examination of modal languages in section 6.5, we tentatively suggested that the accessibility relation on the type of possible worlds ( $W$ ) might qualify as intrinsic structure on  $W$ . At stake is the logicity of the non-S5 modal operators, which cannot be defined with a universal accessibility relation. Now that we have a theoretical characterization of intrinsic structure, it is worth asking whether the accessibility relation qualifies.

The question is whether the accessibility relation must be invoked at the postsemantic level. Suppose we have given a compositional semantics for the language, one that spits out a value in  $(W \Rightarrow V)$  for each sentence, given an interpretation of the language's simple expressions. Two fundamental tasks of the postsemantics are to say what it is for a sentence to be *true* (in terms of these values) and what it is for one sentence to *imply* another. Do either of these tasks require reference to the accessibility relation?

Consider first the account of truth. A sentence with a semantic value  $f \in (W \Rightarrow V)$  is true just in case  $f(w) = \text{True}$ , where  $w$  is the world at which the sentence is uttered (a parameter of the context of utterance). No reference to accessibility is required here. Nor is it required in the account of implication. One sentence implies another just in case there is no interpretation and no context at which the semantic value of the first sentence is True while the semantic value of the second sentence is False.<sup>56</sup> Apparently, then, the postsemantics need not recognize the structure imposed on  $W$  by the accessibility relation. Hence, that structure is not intrinsic to  $W$ , and notions that are sensitive to it (like the S4 necessity operator) are not logical according to the invariance criterion. There is no motivation for allowing logical notions to be sensitive to the accessibility relation on possible worlds and not, say, to the set-theoretic membership relation on  $O$  or the difference between  $A$ 's and

<sup>56</sup>If *frames* are used, they can be construed as features of context (like quantifier domains): see section 6.3.

B's reports on the nine-valued  $V$ . All of these structures are semantically significant only in so far as the language contains expressions that are sensitive to them. Unlike the structure defined by the "logical lattice" on  $V$ , the accessibility relation has no semantic significance at the top level.

### 6.7.5 Application: tense logic

It might be assumed that a similar conclusion would hold for tense logic. For the Prior-Thomason semantics for tense logic in an indeterministic framework (Thomason 1984) is formally similar to the semantics for modal logic. Just as sentences in modal languages are assigned semantic values from the type  $(W \Rightarrow V)$ , where  $W$  is a set of possible worlds, sentences in tensed languages are assigned semantic values from the type  $(MH \Rightarrow V)$ , where  $MH$  is a set of moment/history pairs.<sup>57</sup> And just as modal operators are sensitive to a structure on  $W$  (the accessibility relation), so tense operators are sensitive to a structure on  $MH$  (the structure of temporal succession on moments). Accordingly, one might assume that the structure of temporal succession on  $MH$  should have no more claim to being intrinsic structure than the accessibility relation on  $W$ . If anything, notions sensitive to temporal succession would seem to have *less* claim than notions sensitive to the accessibility relation to being logical, inasmuch as temporal succession is (at least in part) a *physical* idea.

But there is a significant disanalogy between modal accessibility and temporal succession, the upshot of which is that one can make a good case for the logicity (in the sense of invariance) of the basic notions of tense logic. While a *single* world in  $W$  is picked out by a context of utterance, in general only a *range* of values from  $MH$  (the moment-history pairs that have the same moment component) is picked out by context. This difference becomes important in the postsemantic account of *truth* (in the top-level sense appropriate to stand-alone sentences). One cannot say (as we did for modal languages) that a sentence

<sup>57</sup>Moments can be thought of as concrete possible event-locations. Histories are maximal chains of moments, ordered by a relation  $<$  of temporal succession. In an indeterministic framework, it is necessary to work with moment/history pairs rather than simply moments, because a single moment can belong to multiple histories. See Belnap and Green 1994 for a more detailed presentation of the semantics presupposed here.

with semantic value  $f \in (MH \Rightarrow V)$  is true just in case  $f(m/h) = \text{True}$ , where  $m/h$  is the moment/history pair picked out by the context of utterance. For in general, there is no such unique moment/history pair. Nor, it seems, should one say that a sentence with semantic value  $f$  is true just in case  $f(m/h) = \text{True}$  for *every*  $m/h$  in the range picked out by context (as in Dummett 1973:393). That would open up the counterintuitive possibility that (for example) “yesterday morning it was the case that it would rain in the afternoon” could be true today, while “it will rain in the afternoon” was not true yesterday morning. Nor, finally, will it work to posit a “thin red line” marking the one possible history out of the many that pass through a moment that will be actualized. As Belnap and Green 1994 have argued (§6), the “thin red line” picture makes it impossible to give a coherent account of the semantics for *nested* tense operators.

Belnap and Green’s suggestion is more radical than any of these unsatisfactory alternatives. The problem, they argue, is that in the context of an “open future,” there is no top-level unrelativized notion of truth. The connection between the semantic values of sentences and the proprieties for their assertion must be understood in terms of other notions. The pragmatic significance of an assertion, on Belnap and Green’s (tentative) account, is an alteration of normative status comparable to what occurs when one makes a bet. In betting that Bucephalus will win the derby tomorrow, one makes it the case that one is owed money if Bucephalus wins and owes money if he loses. Similarly, in asserting that  $p$  at moment  $m$ , one makes it the case that one is owed credit (for being correct) on the histories on which  $p$  is true at  $m$ , and owed blame (for being incorrect) at the histories on which  $p$  is false at  $m$ . If Belnap and Green are right, then the top-level semantic notion relevant to proprieties of assertion is more complicated than truth *simpliciter*; it is truth parameterized to a structured set of moment/history pairs.

The key word here is “structured.” The top-level proprieties of assertion and inference cannot be articulated without appealing to two structural relations on the category  $MH$ : the relation of *being members of the same history* and the relation of *being temporally*

later than.<sup>58</sup> For the context of assertion will pick out a class C of items of MH: the moment/history pairs that have the moment of utterance as their moment component. One is owed credit or blame only on moment/history pairs that (i) belong to the same history and (ii) are temporally later than some member of C. If it is impossible to specify the pragmatic significance of an assertion without appealing to these structural relations, we ought to take them as intrinsic structure, for just the same reason we regard the distinction between designated and undesignated values as intrinsic. And taking this structure to be intrinsic allows the standard tense-logical operators to count as logical on the invariance criterion.<sup>59</sup>

Thus, despite the apparent limitation of the tense operators to a particular subject matter (tensed discourse), a good case can be made for their logicity. For the structure to which they are sensitive—the fundamental ordering on moments—is essential to a top-level semantic understanding of assertion in an indeterministic world, regardless of the particular expressions our language contains. It is because assertion cannot be understood in abstraction from its temporal setting that the laws of tense logic have a claim to the kind of general applicability characteristic of logic (applicability to thought *as such*).

### 6.7.6 Application: extensional logic

Finally, let us return to the case for which the invariance criterion was originally intended: extensional predicate logic. As we have seen, the introduction of the notion of intrinsic structure—required for the extension of the invariance criterion to sentential and intensional operators, but already motivated by the original criterion’s asymmetrical treatment of V

<sup>58</sup>One can define the relation *are members of the same moment* from these: *a* and *b* are members of the same moment iff they are members of the same history, *a* is not temporally later than *b*, and *b* is not temporally later than *a*.

<sup>59</sup>To see why, reflect that these operators (*Will*, *Was*, *Was-always*, *Will-always-be*) are sensitive only to relations of temporal order between different moment/history pairs, not to any other features of their individual identities. As long as the structure of these order relations is preserved, the underlying set of moment/history pairs can be permuted arbitrarily without affecting the semantic values of the tense-logical constants. [Note added after defense: It is a real question whether there will be any non-trivial permutations of MH that preserve the order relation. If the shape of the “tree” is sufficiently asymmetrical, there may be no way no permute moment/history pairs without spoiling the intrinsic structure. In that case, all sentential operators will come out invariant.]

and  $O$ —reveals a presupposition in the standard application of the invariance criterion to extensional notions. It is always assumed (in effect) that the intrinsic structure of  $O$  is the null structure. Without that assumption, the invariance criterion cannot exclude the set-theoretic membership relation, the mereological sum functor, or even the predicate “is red” as logical constants. But what justifies that assumption?

Now that we have a general characterization of intrinsic structure, we can answer this question. According to the Intrinsic Structure Principle, intrinsic structure on  $O$  is structure that must be mentioned in the postsemantics—generally, in an account of truth and implication. Since these notions can be defined by reference to the semantic values of *sentences* alone, the type  $O$  need not be mentioned at all—nor, *a fortiori*, need any structure on  $O$  be mentioned. Hence it is correct to take the intrinsic structure on  $O$  to be the null structure. Various structures on  $O$  may be *indirectly* relevant to truth and implication—to the extent that the language contains expressions sensitive to them—but these structures are not relevant to truth and implication *as such*, at the top level, independent of any particular expressions of the language. As a result, they are no concern of logic (on the conception being developed here).

But this is only a partial vindication of Sher and the other proponents of the invariance criterion who assume that  $O$  has the null intrinsic structure. For on their telling, it is the invariance criterion *itself* that rules out set-theoretic membership, mereological sum, and so on as logical notions, while in fact, most of the work is being done by an implicit assumption that can only be justified by bringing in considerations quite different from the simple appeal to invariance or “insensitivity to particulars.”

## 6.8 Conclusion

Part of the appeal of the invariance approach to logicality lies in its apparent clarity and simplicity. Rather than mucking around with the difficult issues surrounding the nature of thought or concept use with which issues of logicality have traditionally been connected,

we simply appeal to the *generality* (or 2-formality) of logical notions—that is, their indifference to particular differences between objects—a notion that can be captured in precise mathematical form as permutation invariance (or in the general case, bijection invariance).

In this chapter, I have argued that this putative advantage of the invariance approach is illusory. The appeal to invariance by itself does not yield a determinate criterion for logicality. It appears to do so only because it makes assumptions about the intrinsic structure of the basic semantic types (O and V) that can only be justified by appealing to considerations that are much less clear than the mathematical idea of invariance. It is these assumptions, and not the appeal to invariance itself, that are responsible for excluding the set-theoretic membership relation and the mereological sum operator, among others, from the province of logic.

Yet these assumptions are never argued for—in part because the invariance criterion is usually presented in a way that makes them invisible. Here the functional-categorical presentation of the criterion developed above has two main advantages over other presentations (e.g., Sher’s). The first is that it makes manifest the asymmetry in treatment of the two basic types in extensional logic (O and V). It is as plain as day that O is being permuted and that V is not, and once we see this we cannot help asking for a justification. The second advantage of the functional-categorical presentation is that it can easily be extended to expressions of different grammatical types (sentential operators, functors, etc., in addition to quantifiers and predicates) and to languages with different basic types (for instance, languages with multivalued or intensional sentential operators). These extensions show us that in many cases it is desirable to consider invariance under some, but not all non-trivial permutations of a basic type. I have suggested that in order to accommodate this desideratum, we should think of the basic types as having an *intrinsic structure* that must be respected by permutations. The question then becomes: what structure on a type counts as intrinsic? (Or, alternatively, which structured types are appropriate for use in various semantic endeavors?)

I have defended an answer to this question—the Intrinsic Structure Principle—and ex-

plored its consequences for a number of cases. Supplemented by the Intrinsic Structure Principle, the invariance criterion provides a way to connect questions about logicality with general questions about the pragmatic significance of assertion and inference (and more generally of the *use* of language). But the problem I have identified is independent of my proposed solution. Anyone who appeals to invariance to delineate the logical notions is making assumptions about intrinsic structure on the basic types. If the invariance criterion is to have any claim to be an explanatory criterion for *logicality*, then these assumptions must be justified. Otherwise the criterion is nothing more than a systematization of our antecedent intuitions about which notions are logical; it does nothing to explain *why* these notions are logical. In order to do that, I have suggested, we must engage with more fundamental questions in the philosophy of language: just the kind of questions one might have hoped to avoid by appealing to the tidy mathematical property of permutation invariance.

## Chapter 7

# CONCLUSION

### 7.1 After formality?

At the beginning of *After Virtue*, Alasdair MacIntyre asks us to imagine a future dark ages in which the natural sciences have been lost and only partially recovered. Only fragments of science remain:

... a knowledge of experiments detached from any knowledge of the theoretical context which gave them significance; parts of theories unrelated either to the other bits and pieces of theory which they possess or to experiment; instruments whose use has been forgotten; half-chapters from books, single pages from articles, not always fully legible because torn and charred. (1981:1)

Nonetheless, the people who concern themselves with this lore take themselves to be doing physics, chemistry, and biology, not realizing that "... what they are doing is not natural science in any proper sense at all" (1).

In such a culture men would use expressions such as 'neutrino', 'mass', 'specific gravity', 'atomic weight' in systematic and often interrelated ways which would resemble in lesser or greater degrees the ways in which such expressions had been used in earlier times before scientific knowledge had been so largely lost. But many of the beliefs presupposed by the use of these expressions would have been lost and there would appear to be an element of arbitrariness and even of choice in their application which would appear very surprising to us. What would appear to be rival and competing premises for which no further argument could be given would abound. (1)

In such a world, MacIntyre claims, the language of science would be “in a grave state of disorder” (2). Yet this disorder would be invisible from the perspective of the people who *use* the language. We could see how disordered the conceptual scheme is only by looking at its *history*.

MacIntyre’s thought experiment is preparation for a bold proposal:

...in the actual world which we inhabit the language of morality is in the same state of grave disorder as the language of natural science in the imaginary world which I described. What we possess, if this view is true, are the fragments of a conceptual scheme, parts which now lack the contexts from which their significance derived. (2)

I want to suggest that the language of logical hylomorphism is in a similar state of disorder, one that we can only see clearly by looking at its history. Philosophers still *use* this language in distinguishing the logical from the non-logical. They say that logic is “formal,” that it “abstracts from (or lacks) content,” that it “excludes material considerations.” They use these claims to argue for and against candidate demarcations of logic and to give a significance to projects like logicism and structuralism in the philosophy of mathematics. And they construct technical criteria to make these claims more precise. Yet they lack the system of beliefs that gives the hylomorphic language its point. Hence their language slurs over important distinctions, engenders equivocation, and produces fruitless debates that founder on opposing but equally brute “intuitions” about formality or logicity.

If I am right, then in order to make progress in the philosophy of logic (especially on the demarcation issue), we must look to the *history* of logical hylomorphism, with the aim of understanding and rectifying language that has become “disordered.”

## 7.2 How we got where we are

In chapter 4, I argued that Kant is the source of modern logical hylomorphism.<sup>1</sup> In support of this claim, I offered the following evidence:

- None of Kant’s predecessors think of logic as distinctively formal.
- However, many of his successors recognize Kant as the source of the idea.
- The idea does not appear in Kant’s own works until 1773-5, the period during which Kant first articulates what will become his critical philosophy.
- The texts show Kant defining logic by its *generality* (in my terminology, 1-formality, section 3.1), then *inferring* that it is formal (i.e., 3-formal, section 3.3).
- This inference can be underwritten by an argument that would only have become available to Kant in 1773-5.

By 1773-5, Kant had excellent theoretical reasons for thinking that a general (i.e., 1-formal) logic must also be formal in the sense of abstracting from all semantic content (i.e., 3-formal). His insistence on the formality of logic was never part of a definitional characterization of the subject; it was a substantive thesis (“Kant’s Thesis”) intimately bound up with his deeper philosophical commitments.

Kant’s philosophy of logic had as enduring an influence on later philosophy as his transcendental idealism. By 1837, it has become so common on the continent to characterize logic as “formal” that Bolzano devotes several sections of his *Wissenschaftslehre* to debunking this way of talking. And after Sir William Hamilton’s 1833 *Edinburgh Review* article, many British writers on logic use the hylomorphic terminology.<sup>2</sup> Its Kantian roots largely

<sup>1</sup>I say “modern” because there is a medieval tradition (with sources in antiquity) of distinguishing formal from material consequence. I discuss this tradition in appendix A. I do not believe that this earlier tradition has much to do with the one that starts with Kant; it had nearly died out by the eighteenth century, and Kant never refers to it in any of his discussions of the formality of logic. For more on this, see section 4.2.4, above.

<sup>2</sup>Before Hamilton’s article, no British writer of whom I am aware characterizes logic by its “formality.”

forgotten, “formality” comes to be seen more and more as an essential part of the *definition* of logic, and less and less as a substantive (and potentially contentious) property of it. As it becomes obligatory to find something to mean by the sentence “logic is formal,” “formal” acquires some new meanings. Some philosophers continue to use it in the Kantian sense of 3-formality. But others—some of whom would reject Kant’s claim that logic is 3-formal—use “formal” to mean schematic-formal (section 2.2), 2-formal (section 3.2), or 1-formal—or a confused blend of these. These semantic changes are made possible, in part, by the fact that in Kant, logic is formal in all of these senses, so that it is easy to read one’s favored sense of formality back into Kant’s use of “formal.”

In chapter 5, I showed how Frege gradually but self-consciously rejects the Kantian philosophy of logic, and with it, the hylomorphic terminology. Frege sees clearly that he need not and ought not accept Kant’s Thesis: that he can take logic to be “general,” in Kant’s sense (i.e., 1-formal), without taking it to be “formal,” in Kant’s sense (i.e., 3-formal). He can do this because he rejects the auxiliary premises from which Kant derives Kant’s Thesis. Logic, for Frege, is a substantive science; what distinguishes it from geometry and physics is not that it “abstracts from all content” but that it is normative for thought as such (i.e., 1-formal), not just for thought about a particular domain, such as the physical.

Frege’s clarity could have brought order back to the language of logical hylomorphism, but it didn’t catch on. The logical positivist tradition, influenced by Wittgenstein and neo-Kantianism, takes logic to be 3-formal and appeals to linguistic convention to explain formality. Since the conventions at issue can be syntactically specified, 3-formality begins to be confused with syntactic formality (section 2.1). Meanwhile, “formal” continues to be used to mean 2-formality, 1-formality, and schematic formality. All of these notions are connected *historically*, but they are by no means equivalent. The language of logical hylomorphism is, to use MacIntyre’s words, “in a grave state of disorder.”

I suggest that this history can shed new light on the intractability of contemporary debates about the bounds of logic. These debates can often come down to an unsatisfying battle of “my intuitions against yours” (see section 1.4, above). Perhaps this is because the

antagonists are operating with different conceptions of formality (or topic-neutrality: see section 3.5). The historical analysis I am recommending offers a way to get past the appeal to brute intuitions by *identifying* the different concepts at work and *explaining* how they all came to be thought of as explications of logical formality. Historical self-consciousness can help us avoid appealing blindly to intuitions that come from different strands of the tradition and are held together by historical causes, not (any longer) by reasons.

### 7.3 Applications

Let me offer a few examples of how the conceptual distinctions in chapters 2–3 and the historical analysis in chapters 4–5 can be of use in current debates about the demarcation of logic. Since this is a conclusion, I will be brief and sketchy. Some of these examples are covered in more detail elsewhere in the dissertation; the others should be taken as research proposals.

#### 7.3.1 Topic-neutrality

Since at least Ryle 1954, it has been popular to demarcate logic and the logical notions by their “topic-neutrality” or maximal generality (see section 3.5, above). On its face, this looks like a relatively uncontentious characterization. But it is notoriously difficult to apply: disputes about what counts as logic become disputes about what is topic-neutral. From one point of view, arithmetic and set theory are paradigms of topic-neutrality: they can be applied to virtually any domain, since things of any sort can be counted and collected. From another point of view, however, they have their own special topics: sets and numbers, respectively. From yet a third point of view, there is no such thing as absolute topic-neutrality: topic-neutrality is always a matter of degree.

If we stop at our “intuitions” about topic-neutrality, we are not likely to get far. The solution is to make conceptual distinctions. In section 3.5, I argued that there are three distinct notions of topic-neutrality in play, corresponding to the three notions of formality

(1-, 2-, and 3-formality). Arithmetic and set theory are (arguably) topic-neutral in one sense (1-formal), while they fail to be topic-neutral in another (2-formal). In yet a third sense (3-formal), it is plausible to maintain that no science is absolutely topic-neutral; there is just more and less.

But beyond helping us to *make* these conceptual distinctions, history can help us diagnose the confusion between them. Once we see that our intuitions about formality or topic-neutrality are tracking three different concepts of formality which have come to be confused for *historical* reasons, we can begin to sort and critically evaluate these intuitions as we could not before. We can see, for instance, that historically logic was demarcated by its 1-formality, to which 2-formality and 3-formality were connected (if at all) only by additional philosophical premises. In particular, we can see that 3-formality was taken to characterize logic for reasons deeply bound up with Kant's transcendental idealism (see section 4.4, above). This historical perspective may give us a reason to discount certain of our intuitions about formality or topic-neutrality and to emphasize others.

### 7.3.2 Permutation invariance

Recently, many philosophers and logicians have undertaken to demarcate logical notions by their invariance under all permutations of the domain of objects (see section 3.2 and chapter 6, above). But the notion of formality to which this proposal answers—2-formality—is of little importance historically in the demarcation of logic. Kant did take logic to be 2-formal, but he did not distinguish it in this respect from arithmetic and algebra. And Frege did not take logic to be 2-formal at all. So the key question for the permutation invariance approach to logicity concerns motivation. Why is this an appropriate criterion for *logicity*? The equivocity of “formal” disguises the lack of motivation here. Sher can connect her proposal to the tradition by observing that many philosophers (e.g., Russell and Tarski) have demarcated logic by its formality (cf. 1991:133, 1996:683-4). But it is not clear that there is anything more than a word in common.

The permutation invariance criterion has also been motivated as a way of spelling out

the generality or topic-neutrality of logic (e.g., in McCarthy 1981). But as we have just seen, there are at least three distinct notions of topic-neutrality in play. Why should *this* one be used? Proponents of the permutation invariance criterion may be able to answer this question, but they need to address it.

Finally, it is not clear how 2-formality applies to sentential and intensional operators. In chapter 6, I suggested that the the permutation invariance criterion for logical notions can be extended in a natural way to sentential and intensional operators. However, this extension requires us to distinguish “intrinsic” from “non-intrinsic” structure on basic semantic types, and here 2-formality gives us no guidance. For this purpose, I suggested, we need to invoke 1-formality. The fact that 2-formality (unsupplemented) only yields a criterion for logicity for extensional quantifiers and predicates suggests that it is not sufficiently fundamental to our understanding of logicity.

### 7.3.3 The debate over second-order logic

One of the most intractable debates about the bounds of logic concerns the status of second-order logic. There are two main reasons for taking second-order logic to be non-logical:

- (1) Second-order logic lacks a complete proof procedure (Kneale and Kneale 1962:724, Wagner 1987, Resnik 1988, Jané 1993).
- (2) Second-order logic appears to be committed to an ontology of sets (Quine 1986:66).<sup>3</sup>

Proponents of second-order logic can urge, on the other hand, that

- (3) The semantics and proof theory of second-order quantifiers are natural extensions of the semantics and proof theory of first-order quantifiers (Boolos 1975:514, Shapiro 1991).

---

<sup>3</sup>At any rate, our best explanation of the semantics of the second-order quantifiers requires quantification over subsets of the domain; moreover, whether certain second-order sentences are logically true depends on the truth of substantive set-theoretic claims, like the Continuum Hypothesis (Shapiro 1991:104-5). Boolos 1984 has suggested that English plural quantification can be used to explain the semantics of the second-order quantifiers; if he is right, then this second objection may lapse. See Resnik 1988 for a critique of Boolos’s proposal.

- (4) Unlike first-order logic, second-order logic has sufficient expressive power to characterize mathematical notions like infinity and well-ordering (Boolos 1975:521, Shapiro 1991).

With the exception of the issue of ontological commitment, none of these points are controversial. To make a move in the debate, then, one must argue that the other camp's considerations are irrelevant for the demarcation of logic. But this is tricky: *all* of these considerations seem relevant.

The beginning of wisdom here, I suggest, is to see that our conflicting intuitions may be responsive to competing conceptions of logicity. Consider the question whether second-order logic is formal or topic-neutral. Consideration (2) suggests that it is not; consideration (3) suggests that it is. But perhaps both are right: second-order logic is formal or topic-neutral in some senses (1-formal, 2-formal, schematic-formal) but not in others (3-formal, syntactic-formal). The considerations pushed by both parties in the debate might *all* be relevant to the logicity of second-order logic—on different conceptions of logicity.

If this is right, then in order to make progress, we must distinguish the notions of formality in play and decide which of them ought to be connected with logicity. If all of them had equal claim, then perhaps the correct conclusion would be that there is no such thing as “logic”—or rather, that what was called “logic” turns out to be several distinct disciplines, each properly characterized by one of the (inequivalent) notions of formality. The history suggests, however, that not all of the notions of formality have equal claim to demarcate logic. In chapter 4, I argued that Kant characterizes logic as 3-formal as a *consequence* of his wider philosophical views, and that 3-formality is taken to be *definitional* of logic only later, as the Kantian philosophy of logic becomes entrenched. If this is right, then we should think twice before demanding that our logic “lack substantial content.” Similarly, in chapter 2, I argued that syntactic and schematic formality are not capable of demarcating logic. If this is right, then we should be wary of appealing to them in debating the logicity of second-order logic. Finally, in chapter 5, I showed how Frege could reasonably take his Begriffsschrift to be a logic, despite its non-permutation-invariant

concepts and existential commitments. This ought to make us think twice both about arguing *for* the logicity of second-order logic on the basis of the permutation invariance of the second-order quantifiers, and about arguing *against* the logicity of second-order logic on the basis of its putative existential commitments.

### 7.3.4 The logicity of Hume's Principle

Wright 1983 shows that the Peano postulates for arithmetic can be derived in second-order logic from "Hume's Principle"

**(HP)** the number of Fs = the number of Gs iff there is a one-one mapping from the Fs onto the Gs,<sup>4</sup>

without any appeal to "extensions" (see section 1.2.1, above).<sup>5</sup> Wright's demonstration raises the possibility of a partial vindication of logicism. If (HP) were a principle of *logic*, then Wright's proof would show that all truths of arithmetic are logical truths. Of course, no one takes (HP) to be *logically* true. The most Wright claims is that (HP) is *analytic*. But it is instructive to ask *on what grounds* (HP) is denied to be a logical truth.

It is not sufficient to point out that (HP) contains a primitive functor, "the number of," which is not among the traditional "logical constants." For sometimes it is necessary to broaden the scope of a discipline in order to do the job assigned to it. The theory of relations was not always part of traditional logic, and Aristotle's logic did not even contain rules for the basic sentential connectives. Frege's iterable quantifiers were certainly an innovation. We should leave room for claims that the bounds of logic are wider than was previously thought, provided such claims can be justified on the basis of an antecedently acceptable characterization of logic.

Tradition aside, numbers would seem to have as strong a claim as Frege's extensions to be "logical objects." Frege's arguments that numbers are logical objects are largely

<sup>4</sup>Wright calls this principle ( $N=$ ) (see section 1.2.1, above). The name "Hume's Principle" is due to Boolos, after the citation in FA:§63.

<sup>5</sup>Boolos 1987 has shown that Frege's FA already contains the general lines of the derivation: "[o]nce Hume's principle is proved, *Frege makes no further use of extensions*" (191).

independent of his identification of numbers with extensions: they turn instead on the inferential role of number words (for the “object” part) and the general applicability of arithmetic (for the “logical” part) (see sections 5.2.2 and 5.2.7, above). Thus it would not have been unreasonable for Frege to have reacted to the problems with his theory of extensions by taking “the number of” to be a *primitive* logical functor.<sup>6</sup> Given his antecedent commitment to the logicity of arithmetic,<sup>7</sup> he could have taken his failure to reduce the concept of number to more basic notions as proof that his logic was incomplete and needed to be supplemented by a new primitive.<sup>8</sup> Heck 1997 comes close to considering this proposal: “the question arises why, upon receiving Russell’s famous letter, Frege did not simply drop Axiom V, install Hume’s Principle as an axiom, and claim himself to have established logicism anyway” (274). But Heck, like every other commentator, takes it for granted that (HP) would not be a “principle of logic,” though “perhaps it has a similarly privileged epistemological position.”<sup>9</sup> What is the basis for this universal assumption?

I suggest that the basis for this universal assumption is a shared commitment to

(NE) Logic alone does not imply the existence of any objects.<sup>10</sup>

(NE) is certainly incompatible with the logicity of (HP). And (NE) may well be true. But how is it argued for? Boolos writes:

We firmly believe that the existence of even two objects, let alone infinitely many,

---

<sup>6</sup>The fact that Frege *doesn't* do this does not show that it would have been unreasonable for him to do so.

<sup>7</sup>See especially his claim in FTA that “...we have *no choice* but to acknowledge the purely logical nature of arithmetical modes of inference” (96, emphasis added)—even in advance of actually carrying out the technical reduction (FA:§90 sounds a more cautious note).

<sup>8</sup>Such a primitive would not be *entirely* alien to the logical tradition. *Quantity* has always been a concern of logic, and the traditional quantifiers (“all,” “at least one,” “none”) might be regarded as special cases of numerical quantifiers (“there are exactly two,” “there are more than thirty”). In fact, Boole’s 1868 paper on “numerically definite propositions” has a primitive operator “Nx”, interpreted as “the number of individuals contained in the class x.” In a sketch of a logic of probabilities, Boole argues that “... the idea of Number is not solely confined to Arithmetic, but ... it is an element which may properly be combined with the elements of every system of language which can be employed for the purposes of general reasoning, whatsoever may be the nature of the subject” (1952:166).

<sup>9</sup>Although Bird 1997 argues that (HP) is “broadly logical,” he appears to mean by “broadly logical” what is usually meant by “analytic.”

<sup>10</sup>Of course, standard first-order logic does imply the existence of *one* object, inasmuch as “ $(\exists x)(x=x)$ ” is a theorem. But this is a technical simplification (see Quine 1961:160-2).

cannot be guaranteed by logic alone. After all, logical truth is just *truth no matter what things we may be talking about and no matter what our (nonlogical) words mean*. Since there might be fewer than two items that we happen to be talking about, we cannot take even  $\exists x\exists y(x\neq y)$  to be valid. (1987:199)

Presumably Boolos is not making the (uncontroversial) claim that quantificational logic as we now understand it does not imply the existence of even two objects. For what is at issue is precisely whether our present understanding of quantificational logic is adequate. Instead, Boolos must be making a *conceptual* claim about logic: logical truth is, *by definition*, “truth no matter what things we may be talking about and no matter what our (nonlogical) words mean.”

But if a logic with existential commitments can be ruled out on *conceptual* grounds, just by thinking about what “logic” *means*, then we are going to have trouble making Frege’s Platonist logicism seem so much as coherent. Boolos immediately raises this question:

How then, we might now think, *could* logicism ever have been thought to be a mildly plausible philosophy of mathematics? Is it not obviously demonstrably inadequate? (1987:199-200)

Note that the “inadequacy” to which Boolos is pointing here is independent of the *technical* problems that led Frege to abandon his logicism. Boolos is claiming that Frege’s project can be ruled out *from the start* on the basis of a general characterization of logic. This is surely an intolerable result. In discussing logicism, we ought to use the word “logic” in such a way that it is at least intelligible how someone could have thought that arithmetic (Platonistically construed) could be reduced to logic. In particular, we should not demarcate logic by its 3-formality or 2-formality. If we think that (HP) cannot be logical because we are committed to (NE), then we should *argue for* (NE) on the basis of a conception of logic Frege could have accepted, not one that makes logicism look like a round square.

By demarcating logic by its *1-formality*, we can leave both logicism and (NE) open as conceptual possibilities. There may still be a good argument for (NE): i.e., an argument that the norms for thought as such cannot imply or presuppose the existence of any objects. If so, the argument deserves to be made explicit. The point is not obvious. The rules of chess

presuppose the existence of a board and pieces; might not the norms governing thought *as such* also presuppose the existence of certain objects?<sup>11</sup>

## 7.4 The centrality of 1-formality

One thing chapters 4 and 5 show is that 1-formality is central to how logic is understood in at least one tradition—the one reaching from the neo-Leibnizians through Kant to Frege. Different philosophers in this tradition have different views on the scope of logic and different views about whether logic is 2-formal or 3-formal, but they all agree in demarcating logic by its 1-formality, and that is what allows us to see them as disagreeing about a single subject matter. Nor is this just one tradition of many. It is a particularly important one for those concerned with the demarcation of logic, since it is within this tradition that many of the projects for which the demarcation of logic is important arise (e.g., logicism, structuralism in the philosophy of mathematics: see section 1.2, above).

This suggests that 1-formality ought to have a central place in any answer to the question “what is logic.” However, as I observed at the end of chapter 5, 1-formality has more or less dropped out of twentieth century discussions of the demarcation of logic. There are several reasons for this.

First, the dominant event in twentieth century philosophy of logic—Quine’s refutation of conventionalism—has influenced how people see the options for demarcating logic. There is the Old View, on which logic is distinctively formal (that is, 3-formal); and there is the New View, on which there is no such thing as “Logic” as the tradition conceives it—no philosophically privileged discipline that is different in kind (not just in degree) from the empirical sciences and (perhaps) mathematics. Everything in between these two views gets flattened out. Principled demarcations of logic in general (in the sense of section 1.3.1, above) get conflated with the (failed) conventionalist approaches. And where the choice is between a principled conventionalist demarcation and a pragmatic demarcation,

---

<sup>11</sup>For a view of this kind, see Tennant 1997.

the pragmatic demarcation looks like the only sound alternative.

Second, the appeal to “thought” in 1-formality seems mired in the nineteenth century. Hasn’t philosophy had its “linguistic turn”? Depending on one’s predilections, talk of “thought” is apt to seem either too psychologistic (despite the fact that Kant and Frege were staunch foes of psychologism) or too unscientific to be of much use in demarcating logic.

Third, there are well-known difficulties in making sense of logic as providing “norms for thought.” Harman 1984 argues that

...even the rule “Avoid inconsistency!” has exceptions, if it requires one not to believe things one knows to be jointly inconsistent. On discovering one has inconsistent beliefs, one might not see any easy way to modify one’s beliefs so as to avoid the inconsistency, and one may not have the time or ability to figure out the best response. In that case, one should (at least sometimes) simply acquiesce in the contradiction while trying to keep it fairly isolated. (108)

For example, the set containing all of my beliefs about Cambodia plus the belief that at least one of these beliefs is false is inconsistent, but if there is no way of deciding *which* of my beliefs about Cambodia is false, it may be reasonable to keep them all. It is not clear, then, how the claim that logic provides norms for thought should be understood. Why should we not say that there’s just one norm here—the norm to think what is true—and that logic is simply a very general body of truths?

The fourth difficulty concerns the “as such” in “norms for thought *as such*.” Why should we think that quantifier theory is normative for thought as such, as opposed to thought involving quantifiers, or thought about collections of discrete objects? Suppose there were a language without logical vocabulary (as Brandom 1994 supposes to be possible). In what sense would quantificational or truth-functional logic be normative for thought or reasoning in that language? The universal normative applicability of logic must be based on more than just the ubiquity of the logical constants.

Though I will not try to meet these difficulties here, I do not think they are insoluble. And provided we can solve them, we have good reason to bring 1-formality back to the

center of our thinking about the demarcation of logic. Not only is 1-formality central to the tradition of thinking about logic out of which most of the philosophical projects for which the demarcation of logic matters emerged, it also offers an approach to logic that swings free of the disputed notions of analyticity and *a prioricity*.

In section 6.7, above, I sketched one way in which the thought that logic is 1-formal might shape a technical demarcation of logic. The idea was to use invariance methods to separate notions into two classes:

- notions sensitive to semantic structure required by the particular expressive power of a language (i.e., semantic structure required in order to give a compositional semantics for the language), and
- notions sensitive only to the semantic structure that must be invoked in a general account of the *use* of stand-alone sentences of the language, independent of its particular expressive vocabulary.

I suggested that notions in the second class are logical notions, since the norms governing them depend only on the proprieties for assertion and inference *as such*, and not on the particular expressive power of the language. Truth-functional logic comes to be applicable to reasoning in a language just by virtue of the classification of sentences into true and false for purposes of assessment of stand-alone assertions—a classification that will take place no matter what subject matter is being addressed and no matter what expressive power the language has. Set theory comes into play, by contrast, only if the language actually contains vocabulary sensitive to set-theoretic membership; it is normative only for thought employing the *concept* of set membership. More controversially, non-S5 modal operators come out as non-logical on this criterion, because they are sensitive to a structure on semantic values (the modal accessibility relation) that is required only because the language contains an operator that is sensitive to it—that is, only for the purposes of giving a compositional semantics. In section 6.7.5, I showed how one might use this criterion to argue for the logicity of *tense* operators.

I am not wedded to the chapter 6 approach. There may be other ways of using 1-formality to guide technical demarcation projects. For example, I think that demarcations that appeal to “inferential definitions,” taking logical constants to be those expressions that can be introduced into a language by a conservative set of introduction and elimination rules (e.g., Popper 1947, Kneale 1956, Prawitz 1978, Hacking 1979, Schroeder-Heister 1984, Kremer 1988, Došen 1994), can be profitably conceived as demarcations of logic by its 1-formality.<sup>12</sup> In these approaches, one starts with a set of *structural rules* governing sentences independently of their internal structures (and hence independently of the logical constants within them): rules like transitivity (if A implies B and B implies C, then A implies C) or weakening (if A implies B, then A and C together imply B). These can plausibly be taken to be “norms for thought *as such*,” independent of topic or special vocabulary. Since the logical constants are introduced through introduction and elimination rules that conservatively extend this base of structural rules,<sup>13</sup> they can be conceived as auxiliaries for the study of inferential relations that hold independently of any particular expressions—“punctuation marks,” to use Došen’s term. I will not pursue this line of thought here; I offer it as another way in which the concepts developed in this dissertation might profitably be applied.

## 7.5 Methodological postscript

Many analytic philosophers will find the method of this dissertation somewhat unusual. Though the topic is the philosophy of logic, there are few of the usual hen scratches and numbered theorems. Indeed, much of the work is historical. Yet the aim is not primarily an understanding of the history of philosophy, but an understanding of the various concepts that shape contemporary projects in the philosophy of logic. I have suggested that there is something we can do to get beyond the deadlock of competing intuitions one often finds in contemporary debates about the proper bounds of logic: we can “go historical” and seek to

<sup>12</sup>This is not to say that proponents of these approaches *do* think of them this way: Popper, for one, is clearly aiming at 3-formality.

<sup>13</sup>That is, the new rules for a constant \* do not allow one to prove anything in the \*-free fragment of the language that could not have been proved without them.

understand the *sources* of our intuitions about logicity. This kind of inquiry can help us to articulate and evaluate these intuitions, and (I hope) to make progress where we could not before.

I will attempt no general defense of this method. Those who find the dissertation illuminating will need no advertisement for its method; those who do not will not be convinced by more verbiage. Nor do I intend to disparage work on the philosophy of logic in the classic analytic mold. Far from it: my aim is to provide a general picture within which we can see the significance of various technical projects. What I am urging is that even in the most technical corners of analytic philosophy, research ought to be informed by history. And not just recent history: if we are to fully understand the present situation in the philosophy of logic, we must go back at least as far as Kant.

# APPENDIX

## Appendix A

# THE ORIGINS OF LOGICAL HYLOMORPHISM

Surprisingly, the father of both formal logic and hylomorphism was not the father of logical hylomorphism. Aristotle applies his distinction between form and matter to logic only once: in the *Physics*, he claims that the premises of an inference stand to the conclusion as matter to form (II.3.195a16-21).<sup>1</sup> What he means is that just as stones are required if a house is to be built, so also the premises are required if the conclusion is to be deduced (cf. II.7.198b7-8). There is no trace of a distinction between “formal” and “material” consequence in any of Aristotle’s works. In fact, the concepts of *form* and *matter* are entirely absent from the *Organon*.<sup>2</sup>

Who first used the concepts of matter and form to characterize logic, and why? The first *extant* use of the hylomorphic terminology in connection with logic is in Alexander of Aphrodisias’s commentaries on Aristotle’s logic works (c. 200 A.D.; Lee 1984:39),<sup>3</sup> but as Ebbesen 1981 notes, “Alexander does not give the impression that he is using a terminology

---

<sup>1</sup>For the uniqueness claim, see Barnes 1990:40.

<sup>2</sup>Burnyeat A:3. Burnyeat argues that in *Metaphysics Z*, Aristotle sharply separates between “abstract, logical” (*logikōs*) discussions and discussions that bring in “principles appropriate to the subject,” i.e., the *physical* concepts of matter and form. This suggests that it is no accident that Aristotle refrains from applying his hylomorphic concepts to logic.

<sup>3</sup>At the end of this chapter, I give approximate dates for all of the ancient and medieval writers discussed herein.

of his own invention” (95; cf. Barnes 1990:42-3). This use of the terminology is almost certainly the invention of an earlier Peripatetic, but beyond that we know virtually nothing. Nor do we have any evidence that would help us to determine *why* hylomorphic terminology was applied to logic (Barnes 1990:43). Alexander and the other Hellenistic commentators who use the terminology do not say much to explain its significance or usefulness. They just use it as if it is familiar.<sup>4</sup>

The situation is much the same in fourteenth century scholastic logic. The distinction between logical matter and logical form is articulated with great precision, but never justified or motivated. Buridan’s discussion is representative:<sup>5</sup>

I say that in a proposition (as we’re speaking here of matter and form), we understand by the “matter” of the proposition or consequence the purely categorical terms, i.e. subjects and predicates, omitting the syncategorematic terms that enclose them and through which they are conjoined or negated or distributed or forced to a certain mode of supposition. All the rest, we say, pertains to the form. (TC:I.7.2)<sup>6</sup>

Accordingly, a formal consequence is one in which no categorematic terms occur essentially—one that remains valid no matter what the matter, provided we keep the form the same—while a material consequence is one that fails to hold “in all terms (keeping the form the same)” (TC:I.4.2-3).<sup>7</sup> What we do not find in Buridan is a discussion of the *point* of these distinctions. What is the *significance* of the distinction between logical form and

---

<sup>4</sup>The Stoics talk of *forms* or *patterns* (*schēmata*) of argument and recognize “bad form” as one source of invalidity in arguments (Sextus Empiricus *PH* II.146). But this shows only that they think of their logic as *syntactic*- or *schematic-formal* (in the sense of chapter 2, above), not that they accept any form of logical hylomorphism. (Aristotle talks of *schēmata*, too: his syllogistic “figures.”) As far as we know, the Stoics never contrast the form and matter of arguments in the way that later Peripatetics do.

<sup>5</sup>Unless otherwise indicated, all translations from Greek or Latin authors are my own.

<sup>6</sup>Buridan is even explicit about what features are included in “all the rest,” listing copulas, negations, signs of quantity, number of sentences and terms, order of signs, relations of relative terms, and modes of signification (I.7.2). Given the presence of the last item, Buridan’s distinction between matter and form is not purely syntactic, but it is as close as one could get to purely syntactic without moving to a language more regimented than Scholastic Latin. Without it there would be no way to distinguish the forms of “Socrates runs (*Sortes currit*)” and “a man runs (*homo currit*).”

<sup>7</sup>Similar definitions can be found in Pseudo-Scotus and Albert of Saxony. In England, the distinction between formal and material consequence is drawn differently (Ockham, Burley), but there is another distinction (*gratia formae/gratia materiae*) that corresponds to Buridan’s. See section A.5, below.

logical matter, or between formal and material consequence? Why are these distinctions drawn in the way they are? What philosophical purpose do they serve? It is not clear, for instance, where Buridan's notion of formality falls in the categories articulated in chapter 3: is formality a matter of abstraction from the content of thought (3-formality), indifference to the particular identities of objects (2-formality), or constitutive normativity for thought as such (1-formality)—or none of these? As with Alexander, the problem is that the scholastics appear to be handing down a distinction that is already well established.

In this case, however, we can locate a probable antecedent: Abelard's twelfth century distinction between *perfect* inferences, which are valid in virtue of their structure (*complexio*), and *imperfect* inferences, which take their validity from "the nature of things" (*natura rerum*). Though he does not use the hylomorphic terminology, Abelard is clearly articulating a version of 3-formality. And unlike later medievals, Abelard offers a vigorous defense of his position against an opponent who would claim that *all* valid inferences take their validity from "the nature of things." There is good reason to think that the Abelardian distinction is an ancestor of the fourteenth century distinction between formal and material consequence (Kneale and Kneale 1962:274-5, 279; cf. Stump 1989:127, Green-Pedersen 1984:198). But even apart from this connection, Abelard's arguments are illuminating. They are sensitive to the difference between 3-formality and schematic formality, and they make a good case for the claim that categorical syllogisms are 3-formally valid. As I will show, however, Abelard's case depends on some characteristically medieval assumptions—assumptions we no longer find plausible.

The heart of this chapter is an examination of Abelard's arguments for the distinction between perfect and imperfect *inferentia* (§A.4). Before looking at these arguments, however, it will be useful to see how the dialectical framework in which they take place evolves out of Hellenistic debates over the status of logic.

## A.1 Aristotle and the commentary tradition

### A.1.1 Aristotle and formal logic

There are good reasons to acknowledge Aristotle's categorical syllogistic as the first formal logic. It employs schematic letters for generality, so it is schematic-formal. But it is also presented as being normative for reasoning about any subject matter whatsoever, no matter what the epistemic status of its premises—that is, as 1-formal. Its claim to 1-formality rests on the general account of propositional structure on which it is based: any claim about the world, Aristotle thinks, must have one of the four categorical forms and hence be subject to the norms of categorical syllogistic.<sup>8</sup>

In all these respects, the theory of categorical syllogistic differs fundamentally from Aristotle's earlier vision of the syllogism. It is now generally accepted that "syllogism" does not *mean* "categorical syllogism" in Aristotle, since the *Topics* and *Sophistical Refutations*, which also concern themselves with syllogisms, must have been written before Aristotle developed the categorical syllogistic (Allen 1995:177-9). Nor does it mean merely "deductively valid argument" (Frede 1974:115, Barnes et al. 1991:21). A syllogism is "... an argument in which, some things having been set down, something other than the things laid down follows by necessity *through the things laid down*" (*Top* 100a25-27)—or, in the *Prior Analytics* version, "*through their being so*" (*AnPr* 24b18-20). The force of the emphasized phrase, which Aristotle glosses as "not needing any term from without (*exōthen*) in order for the necessity to come about" (*AnPr* 24b21-2), is to rule out arguments in which the necessity of consequence depends on unstated assumptions. A syllogism, then, is a deductively valid argument in which all of the assumptions on which the necessity of consequence depends have been made explicit (Frede 1974:115). So far, the *Topics* and the *Prior Analytics* are in

---

<sup>8</sup>Aristotle's logic is not *syntactic-formal*: as Lukasiewicz 1957 says, it is "... formal without being formalistic" (15; cf. Mueller 1974:51). The Peripatetics could never understand the Stoic insistence on the importance of syntactic rules (Frede 1974:108-110). Alexander rejects Stoic syllogisms like "if it is day, it is day; it is day; therefore, it is day" on the grounds that "... the shape (*schēma*) of the expression is not sufficient to make a syllogism, but it is necessary first that what is signified by the expression be capable of proving something" (*Top*:19-21).

agreement. The crucial difference is that where the *Prior Analytics* offers a *general* theory of the syllogism, applicable to demonstrative, dialectical, and eristic arguments alike, the *Topics* offers a theory of the *dialectical* syllogism: a set of heuristics for constructing arguments from reputable premises, in the absence of specialized scientific knowledge. As James Allen 1995 has argued, it is a mistake to assume that the theory of the *Topics* is intended to have *any* application to *demonstrative* syllogizing, even though some of the topical maxims resemble formal logical laws (for example, the maxim of *Top* 111b18-21 is essentially *modus ponens*). Demonstrative reasoning relies not on *Topics* but on the special first principles of specific sciences (*Rhet* II.2.1358a2-35; for interpretation, see Allen 1995:192-9). In short, prior to the invention of the categorical syllogism, Aristotle has no account at all of the “common logical form” of dialectical and demonstrative syllogisms. The *Prior Analytics* offers the first theory of the syllogism, *as such*.

The theory of the syllogism presupposed in the *Topics* and the earlier strata of the *Rhetoric* cannot be said to be formal in any of the senses in which the categorical syllogistic can be. It is not 1-formal, because it does not provide any common norms for syllogisms in the demonstrative sciences: rather, the principles of the particular sciences are themselves regarded as the standards for correct demonstrative syllogism. Even the norms it offers for *dialectical* syllogisms are not completely general. Although the topics include general rules or maxims—for example, “if contrary attributes belong to the genus, they also belong to the species” (cf. *Top* II.4.111a14-5)—Aristotle frequently notes that these rules have exceptions (objections or *enstaseis*: 115b14, 117a18, 117b14, 121b30, 123b17, 124b19, and 128b6, just to name a few). Thus, even considered as norms for dialectical syllogisms, topical maxims are not formal—not even schematic-formal.

Before Aristotle’s invention of the categorical syllogistic and his use of it as an account of the correctness of syllogisms *as such*, then, logical hylomorphism would not have been possible. Allen 1995 is right to credit to Aristotle the “insight...that the validity of an argument is due to its form, not its content, and that this form can be isolated and made the object of systematic study” (191). However, it is important to recall that Aristotle did

not himself characterize his insight in these terms. It was left to later Peripatetics to take that step.

### A.1.2 Logical hylomorphism in the Greek commentaries

At the beginning of his commentary on the *Prior Analytics*, Alexander of Aphrodisias compares the syllogistic figures (*schēmata*) to a “common matrix” (*tupos koinos*) in which many different matters can be molded into the same form (*eidōs*) (*AnPr* 6.16-21). This division of syllogisms into form and matter is echoed in later commentators, such as the neo-Platonist commentator Ammonius:<sup>9</sup>

In every syllogism there is something analogous to matter and something analogous to form. Analogous to matter are the objects (*pragmata*) themselves by way of which the syllogism is combined, and analogous to form are the figures (*schēmata*). (*AnPr* 4.9-11, trans. Barnes 1990:41)

The point of the schematic letters in Aristotle’s exposition is, Alexander says,

... to indicate to us that the conclusions do not depend on the matter (*ou para tēn hulēn*) but on the figure (*para to schēma*), on the conjunction of the premisses (*tēn toiautēn tōn protaseōn sumplokēn*), and on the modes (*ton tropon*). For so-and-so is deduced syllogistically not because the matter is of such-and-such a kind but because the combination (*suzugia*) is so-and-so. The letters, then, show that the conclusion will be such-and-such universally, always, and for every assumption. (*AnPr* 53.28-54.2, trans. Barnes et al. 1991)

The validity of *non*-syllogistic arguments, on the other hand, depends on the particular matter (52.19-25). Among these are syllogisms in the second figure with two affirmative premises in which the terms are all necessarily coextensive (definitions or propria, 344.28-31), for example:

All humans are mortal rational animals.  
All humans are featherless bipeds.  
Therefore, all mortal rational animals are featherless bipeds.

---

<sup>9</sup>Starting with Porphyry, the neo-Platonic commentators accept and expound Aristotelian logic, provided it is not taken as a key to ontology (which must be done in strictly Platonic terms) (Ebbesen 1982:103, 1981:134-9). By this time, Stoic logic had virtually disappeared, except for the traces it left in the scholastic tradition (Ebbesen 1982:103).

Here the conclusion follows necessarily from the premises, not on account of the (invalid) syllogistic form, but because of the peculiarity of the matter.<sup>10</sup>

Both Alexander and Ammonius see the theory of categorical syllogistic as an account of the syllogism *as such* (*haplōs*); that is, of the form (*eidōs*) common to demonstrative, dialectical, and eristical syllogisms (Alex. *Top* 2.1-3.4; Ammon. *AnPr* 4.1-7). Alexander observes that the following two syllogisms do not differ from one another in respect of their form (*kata to eidōs*), although (due to the difference in their matter) the first is demonstrative and the second dialectical (*Top* 2.26-3.4; cf. Flannery 1995:120-1):

Pleasure is incomplete.	That which is good produces good things.
Nothing good is incomplete.	Pleasure does not produce good things.
Therefore, pleasure is not good.	Therefore, pleasure is not good.

The hylomorphic terminology is also deployed to describe the difference between the two sorts of fallacious arguments: materially defective arguments (*hēmartēmenos para tēn hulēn*, i.e., having a false premise) and formally defective (*hēmartēmenos para to eidōs*, i.e., invalid) ones (Alex. *Top* 20-1; cf. the other passages cited by Ebbesen 1981:95-6).<sup>11</sup>

This evidence suggests that logical hylomorphism was fairly widespread among the Greek scholastics.<sup>12</sup> However, it is too meager to yield confident answers to the questions in which we are most interested. Why do these commentators bring the hylomorphic terminology to

<sup>10</sup>A similar distinction is made among conversions of categorical propositions, which do not count as syllogisms because they have only one premise (see Barnes 1990:625). For example, universal affirmatives convert to universal affirmatives only in particular material instances (*epi hulēs*, 35.3-4) : e.g., “all humans are mortal rational animals” converts to “all mortal rational animals are human.” But universal affirmatives always convert to particular affirmatives: such conversions “do not depend... on the peculiarities of matter (which is different in different cases) but on the nature of the figures themselves” (35.6-9, trans. Barnes et al. 1991).

<sup>11</sup>Ebbesen notes that this distinction is probably an Aristotelianizing of the Stoic distinction between two kinds of “false” arguments: those with a false premise and those that cannot be reduced to the five indemonstrables (Ebbesen 1982:125). Commentators on Aristotle use it to illuminate Aristotle’s division of eristical syllogisms into (a) real syllogisms from premises that are merely apparently reputable and (b) merely apparent syllogisms (*Top* I.1.100b23-6). The distinction between material and formal defect reappears in Michael of Ephesus’s commentary on the *Sophistical Refutations* (*SE* 4.9-13, 4.29-31, 6.31-7.1, 7.9-11, 135.26-7), which was translated into Latin by James of Venice in 1130 (Ebbesen 1982:108; for the identification of Pseudo-Alexander with Michael, see Ebbesen 1981:268).

<sup>12</sup>For general discussions of logical hylomorphism in the commentators, see Lee 1984:37-44, Barnes 1990, and Flannery 1995:ch. 3.

bear on logic? Does their use of the concepts of form and matter tell us anything about how they think of logic? Or is it just an unreflective extension of familiar Aristotelian machinery to an analogous case? What is the *significance* of saying that the validity of the syllogism depends not on the matter, but on the form?<sup>13</sup>

It is clear that Alexander intends something more than *schematic* formality, because he does not think that all arguments that can be presented using schematic letters are syllogistic. In particular, the conclusion of the argument

(GT) A is greater than B.  
 B is greater than C.  
 Therefore, A is greater than C.

follows from the premises “on account of the peculiarity of the matter” (*para tēn tēs hulēs idiotēta*, *AnPr* 344.28-9). In order to see why Alexander does not count this argument as formally valid, we need to understand why he does not count “is greater than” as a feature of the argument’s form (like “belongs to all”).<sup>14</sup>

Although he never gives a general criterion for what we would call the “logical constants,” Alexander does wonder why *modal* vocabulary is not part of the matter of a syllogism:

For the fact that a predicate belongs in this way rather than in that way is a material difference. Differences of this sort among propositions will seem to bear not on an argument’s being a syllogism *simpliciter* (*haplōs*) but on its being this or that kind of syllogism—demonstrative, say, or dialectical. (*AnPr* 27.29-28.2, trans. Barnes et al. 1991)

His answer is that consideration of the modal vocabulary is “useful” and indeed “necessary” for the methodical study of syllogistic (28.2-4, 18). Propositions convert differently

<sup>13</sup>Frede 1974 claims that neither Stoics nor Peripatetics ever say that an argument is valid because of its logical form (103). He rightly points out that to say that an argument form holds in all matter is not yet to say that its instances are valid *because of* their form. However, the evidence I have cited above suggests that the commentators would make the latter claim as well as the former. For syllogisms are *contrasted* with arguments whose validity depends on their matter, and although syllogisms are not said literally to be valid in virtue of their *forms*, they *are* said to be valid in virtue of their *schēma*, *sumplōkē*, or *suzugia*, which surely amounts to the same thing (see Barnes 1990:40 n. 62).

<sup>14</sup>See Łukasiewicz 1957:14-15.

depending on their modes, and so the modes of the premises and conclusion must be taken into account in determining which syllogisms are valid (28.4-13). *Because* the modes must be taken into account in order to give a systematic account of valid syllogisms, they must be “annexed” (“apart from the matter,” *chōris tēs hulēs*) into the form of the propositions (28.13-14). On the other hand, modal words that “do not bear on the generation or differentiation of syllogisms”—for example, “badly,” “quickly,” or “concisely”—are presumably part of the matter (28.20-4, trans. Barnes et al. 1991).<sup>15</sup>

But this criterion is not much use for deciding whether “greater than” should count as part of the form of the argument (GT) or as part of the matter. For what is at stake is precisely whether (GT) *is* a syllogism. In order to determine whether “greater than” bears on the “generation or differentiation of syllogisms,” we have to decide whether (GT) is a syllogism; but in order to decide whether (GT) is a syllogism, we need to know whether “greater than” counts as matter or form. Alexander injects content into the emptiness of this circle by appealing to a characteristic feature of *categorical* syllogisms: all must have a universal premise (344.23-5, 345.19-20). But this move just begs the question.<sup>16</sup>

Alexander’s view seems to have been that only categorical syllogisms (and the conversions of categorical propositions) are valid in virtue of their forms. Although other inferences may be valid—in the sense that their conclusions necessarily follow from their premises—their validity is based on peculiarities of their “matter” or terms. But Alexander gives no clear account of just what it is that non-categorical arguments lack. As a result, it is not clear whether the formality he attributes to categorical syllogisms is 1-formality, 2-formality, or 3-formality—or something quite different.

In the later, neo-Platonist commentators, there is a hint of the view that logic is 3-formal. Ammonius attempts to reconcile the Stoic view that logic is a part of philosophy with the Aristotelian view that logic is not a part of philosophy but merely an *organon*

<sup>15</sup>For a contrasting interpretation of this passage, Flannery 1995:111-4. The passage is also discussed in Barnes 1990:47-53.

<sup>16</sup>Barnes 1990 suggests, on slim evidence, that the Peripatetics distinguish the logical constants on the basis of their universal applicability or “topic-neutrality” (52-3). I am inclined to think, however, that Alexander had no such general criterion in mind.

(tool). Logic, he says, can be considered in two ways:

If you consider the arguments together with the things (*pragmatōn*), e.g. the syllogisms themselves together with the things that are their subject matter (*meta tōn pragmatōn tōn hupokeimenōn autois*), [logic is] a part [of philosophy]. But if you consider the empty rules apart from the things (*psilous tous kanonas aneu tōn pragmatōn*), it is an *organon*. (*AnPr* 10.38-11.3)<sup>17</sup>

The terminology seems to come from Plotinus, who contrasts Platonic dialectic, which is a part of philosophy and concerns itself with real beings and things (*onta, pragmata*), with Aristotelian logic, which provides “empty theorems and rules” (*psila theōrēmata kai kanones*) (*Enn* I.iii.5.10-12, cf. 4.18-20).<sup>18</sup> To say that the rules are “empty” is apparently to say that they abstract entirely from relation to real things—that is, that they are 3-formal. Plotinus’s pupil Porphyry (a key influence on all later scholasticism, Greek and Latin) seems to have exploited this view to argue that there is no real incompatibility between Aristotelian logic and Platonic metaphysics (Ebbesen 1981:134-6), since logic says nothing about ontology. Again, however, our evidence is slight. Only Porphyry’s *Introduction to the Categories* and his commentary on the *Categories* survived into the middle ages.

### A.1.3 Logical hylomorphism in Boethius

It is likely that logical hylomorphism passed into the Latin medieval tradition through Boethius. Boethius’ influence on Abelard and his contemporaries was, as we will see, immense. Until the Latin West’s rediscovery of Aristotle’s *Analytics* and *Topics* in the twelfth century, Boethius’ treatises were the main source of knowledge about the syllogism (categorical and hypothetical) (Ebbesen 1982:105, 122).

Barnes 1990 suggests that much of Boethius’ terminology is meant to translate the logical-hylomorphic vocabulary in the Greek commentators:

In his *de hypotheticis syllogismis* he invokes the Peripatetic distinction in a variety of ways: *propositionis ipsius conditio* contrasts with *rerum natura* (II

<sup>17</sup>Philoponus expresses a similar view, but with “universal rule” (*katholikos kanōn*) instead of “empty rule” (*AnPr* 47.1-5). See Lee 1984:42-4, Flannery 1995:117.

<sup>18</sup>I owe the reference to Flannery 1995:117 n. 29.

ii 4); *propositionum complexio* with *rerum natura* (II ii 5); *complexionis natura* or *figura* with *termini* (II iii 6; iv 2); *complexionis natura* with *terminorum proprietates* (II iv 3; x 7; III vi 5); *complexio* with *termini* (II xi 1). (42)

Barnes' claim is plausible: *complexio* is cognate to the Greek *sumplokē* (Alex. *AnPr* 53.30; Philop. *AnPr* 321.3, 349.30), *figura* translates *schēma*, and *rerum natura* corresponds to *phusis tōn pragmatōn* (Ammon. *in Int.* 113.13-15, quoted in Barnes 1990:44). As we will see, the Boethian terminology turns up again in Abelard, who distinguishes between consequences that take their truth from their construction (*complexione*) and those that take their truth "from the nature of things" (*natura rerum*). Boethius is therefore a bridge between the Greek scholastic tradition and the later Latin one.<sup>19</sup>

However, the distinction Boethius draws between hypothetical syllogisms that are valid by virtue of "the construction of their propositions" and those that are valid by virtue of "the nature of the things, in which alone these propositions can be asserted" (*DHS* II.ii.4-5) is not the same as the distinction Alexander draws between syllogisms valid on account of their matter and syllogisms valid on account of their form. Nor is it the same as Abelard's distinction between perfect and imperfect inference. The syllogisms Boethius calls valid "according to terms" are cases of affirming the consequent or denying the antecedent that happen to be valid in all substitution instances *for which the major premise is true*. For example, in the inference "if it is not *a*, it is *b*; but it is *a*; thus it is not *b*," Boethius claims, the major premise can only be true when the terms *a* and *b* are *contraries*, like "day" and "night." But when *a* and *b* are contraries, and it is *a*, then it follows that it is not *b*. Hence the inference is valid for all substitution instances in which the premises are true. Where Abelard and Alexander are pointing to a class of inferences that are valid *despite* having formal counterexamples, the syllogisms Boethius calls valid by "the nature

<sup>19</sup>The connection is urged by Green-Pedersen 1984:198. Note that even Abelard's word "perfect" appears in Boethius (*DHS* II.ii.6). There is another relevant passage at *ICT* 1046, where Boethius distinguishes between the "matter" and "form" of arguments, comparing these to the stones and their arrangement in a wall. To contemplate the matter of an argument is to examine "the nature of the propositions themselves—whether they are true and necessary, whether they are verisimilar, or whether they are used in sophistries." To contemplate the form is to consider "the junctures and composition of the propositions among themselves" (trans. Stump 1988).

of the things, in which alone these propositions can be asserted” are counterexample-free. They are distinguished from syllogisms valid by “the construction of their propositions” not because they have counterexamples, but because their freedom from counterexamples depends on peculiarities of the matter “in which alone” they can be instantiated, not on their “construction.”

In fact, Abelard thinks Boethius’ claim that “if it is not *a*, it is *b*” can only be true when *a* and *b* are contraries is simply *wrong*: “if it is not *a*, it is *b*” can be true, he notes, when *a* is “animal” and *b* is “non-man” (*D* 499). Abelard agrees that some cases of affirming the consequent are valid “due to the terms” (501-2), but he does not mean the same thing by this phrase as Boethius.

Thus, although Abelard is probably taking over Boethius’ *vocabulary*, he is not taking over Boethius’ *distinction*. Nor (as far as we know) did he have any other source for the distinction articulated by Alexander and the other Greek scholastics. In order to understand the background of Abelard’s account of the special character of the syllogism, we must look at two other developments in later antiquity which made the issue of the distinctive character of syllogistic more pressing. The first is the debate over “unmethodical arguments”: valid arguments that cannot readily be fit into the framework of either the (Peripatetic) categorical or the (Stoic) hypothetical syllogistic. The second is the gradual transformation of the Topics from heuristics for argument discovery into inferential norms.

## A.2 Unmethodically conclusive arguments

For both Stoics and Peripatetics, the point of logical theory is not to codify necessary inferences, but to provide a standard for the *explicitation* of inferences. Thus both schools recognize inferences that are conclusive (i.e., in which the conclusion follows necessarily from the premise) but not *syllogisms*. The Aristotelian definition of syllogism requires not only that the conclusion follow necessarily from the premises, but that it follow “through their being so” (*tōi tauta einai*, Arist. *AnPr* I.1.24b20). Aristotle glosses this phrase as “not

needing any term from without (*exōthen*) in order for the necessity to come about” (24b21-2): the point is apparently to rule out necessary inferences that depend on unexpressed premises (at least that is how the later Peripatetics understand it, Alex. *AnPr* 21, 344). Similarly, the Stoics divide conclusive (*perantikoi*) arguments—i.e., arguments in which the premises are incompatible with the negation of the conclusion—into two classes: *sylogistic* (*sullogistikoi*) and merely conclusive (*perantikoi eidikōs*) (Diogenes Laertius *LP* VII.77-8). What distinguishes the sylogistic arguments from the merely conclusive ones is that they can be reduced (using second-order proof reduction rules, *themata*) to the five canonical forms of arguments the Stoics call “indemonstrable” (VII.78). Again, the effect is to rule out arguments that are conclusive in virtue of some unexpressed premise.<sup>20</sup>

For both Peripatetics and Stoics, then, syllogisms are not merely necessary inferences, but inferences in which all the assumptions on which the necessary consequence depends have been made explicit.<sup>21</sup> The role of sylogistic theory is to provide a framework for this complete explicitation. Of course, Peripatetics and Stoics find completely different frameworks appropriate for this purpose: the Peripatetic paradigm of a completely explicit inference is the categorical syllogism, whereas the Stoic paradigm is the hypothetical syllogism. Oddly, neither side recognizes the other’s syllogisms as syllogisms (Frede 1974:100; cf. Mueller 1960). The Stoics and Peripatetics regard term logic and propositional logic, which we see as complementary, as conflicting doctrines.

<sup>20</sup>Frede 1974 suggests that “...the Stoics thought that every valid argument will turn out to be a syllogism if only we supply the premises which have been taken for granted and formulate them properly...” (103; cf. Mueller 1960:179-80). Against this, Barnes 1990 argues that the Stoics may have thought that at least some conclusive arguments—the “unmethodically conclusive arguments,” discussed below—are “untreatable by the science of logic” (hence the name “*unmethodically conclusive*,” 81). The evidence is probably too meager to decide either way.

<sup>21</sup>This much is common to the Stoic and Peripatetic understanding of syllogism. There are also some differences: the Peripatetics seem to have thought of syllogisms as *arguments*, not just inferences, and built into the notion of syllogism the requirement that something *new* be proved (Frede 1974:117), while the Stoics required that syllogisms be in a canonical (linguistic) form, as the Peripatetics did not (Frede 1974:102-3, Alex. *AnPr* 373.29-35; see note 8, above). These differences will not concern us in what follows.

### A.2.1 The debate over unmethodically conclusive arguments

This conflict comes out most clearly in Alexander’s debate with opponents he calls “the moderns” (*neoteroi*, *AnPr* 390.16-18)—most likely late Stoics.<sup>22</sup> Both Alexander and the *neoteroi* agree that the arguments the Stoics call “unmethodically conclusive”—for example,

(EQ)    A is equal to B. B is equal to C. Therefore, A is equal to C.	(GT)    A is greater than B. B is greater than C. Therefore, A is greater than C.
---	---

—are not syllogisms, because their necessity depends on an additional assumption beyond the stated premises (*AnPr* 22.4-7, 345.13-14, 24-7).<sup>23</sup> They are, as the Peripatetics would put it, valid because of their matter, not their form.<sup>24</sup> Alexander goes through great lengths to persuade us that such inferences can be put into the form of a categorical syllogism by conjoining the premises and adding another, universal premise (344.15-20).<sup>25</sup> On the Stoic side, Posidonius seems to have argued that arguments like (EQ) are valid in virtue of an implicit axiom (*sunaktikoi kata dunamin axiōmatos*,<sup>26</sup> Galen *IL* XVIII.8, Kidd 1978:279, Ebbesen 1981:113). It is common ground between Stoics and Peripatetics that such arguments are not syllogisms.

The disagreement concerns the relation between such uncontroversially *non*-syllogistic arguments and the arguments each party takes to be genuinely syllogistic. The *neoteroi* argue that the Peripatetics’ categorical syllogisms are like unmethodically conclusive argu-

<sup>22</sup>For discussions of the identity of the *neoteroi*, see Mueller 1974:58-9, Barnes 1990:74-5, and Kieffer 1964:130-3. Ebbesen’s suggestion of Posidonius (1981:113) is attractive but speculative.

<sup>23</sup>(EQ) comes from Euclid’s first proposition (*E* I.1; Alex. *AnPr* 20.4-5); (GT) can be found in *E* I.18. For other examples of unmethodically conclusive arguments, see *AnPr* 344.9-346.6, Pseudo-Ammonius *AnPr* 70.10-13, and (GT) above. Discussions of unmethodically conclusive arguments can be found in Frede 1974:102, Mueller 1960, 1974:59-66, Ebbesen 1981:112-4, Barnes 1990:§IV.

<sup>24</sup>Alexander says that the conclusions of inferences like (GT) and (EQ) follow from the premises “asylogistically and through the peculiarity of the matter of the [premises] laid down” (344.28-9). And Pseudo-Ammonius writes: “Thus let not the geometers say: ‘Since A is equal to B and B is equal to C, therefore A is equal to C.’ For they deduce truths not because of the combination (*plokē*) but because of the matter. That is why the Stoics call them unmethodically concluding” (*AnPr* 70.10-13, trans. Barnes 1990:80).

<sup>25</sup>The claim is logically naïve—see Mueller 1960:176, 1974:42, Barnes 1990:101-4—but that is not our present concern.

<sup>26</sup>Posidonius appears to be using *axiōma* with its Peripatetic sense of “self-evident proposition,” not its Stoic sense of “proposition” (Mueller 1974:64).

ments, and hence not syllogisms at all (Alex. *AnPr* 345.15-18): only hypothetical syllogisms are really syllogisms (260.28-9, 262.28-9). Alexander argues for the converse position: that it is Stoic *hypothetical* syllogisms that are like unmethodically conclusive arguments (i.e., conclusive, but not syllogistic—390.1619, 348.31-2), and that only categorical syllogisms are truly syllogistic. In effect, each party to the debate argues that the inferences the other party considers *formal* are in fact *material* and depend on unstated assumptions, just like unmethodically conclusive arguments.

### A.2.2 The neoterioi's case

Alexander does not say *why* the *neoterioi* think that categorical syllogisms are like unmethodically conclusive arguments; he is interested in rebutting the view, not sympathetically expounding it. We are left to speculate. As we have seen, Posidonius probably argued that (EQ), (GT), and similar arguments get their force from an implied axiom. And in fact, Euclid's first Common Notion ("Things which are equal to the same thing are also equal to one another," *E* 155; cf. Kidd 1978), which Proclus calls an "axiom" in his commentary (cited by Heath in Euclid *E* 121; cf. Ross 1949:510-11), is just what is needed to validate (EQ). It is possible that further reflection on the axiomatic basis for the "unmethodically conclusive" arguments used in mathematics led Posidonius to ask whether categorical syllogisms, too, depend on an implied axiom for their cogency (so speculates Ebbesen 1981:113). Galen tells us that Boethus of Sidon, a Peripatetic, allowed that some hypothetical syllogisms were prior (in some sense) to categorical syllogisms (*IL* VII.2, Kieffer 1964 *ad loc.*, Ebbesen 1981:113, Frede 1974:123). Perhaps Peripatetics began to reflect on the principles of propositional logic Aristotle used in developing syllogistic theory (especially in arguments by *reductio* for the validity of some of the syllogistic moods).

Another tempting conjecture (for which, again, there is no direct evidence) is that the *neoterioi* noticed the formal similarity between arguments like (EQ) or (GT) and categorical syllogisms in Barbara (Bar), phrased in Aristotle's usual manner:

(GT)    A is greater than B.                    (Bar)    A belongs to all B.

B is greater than C.	B belongs to all C.
Therefore, A is greater than C.	Therefore, A belongs to all C.

If (GT) depends for its validity on an assumption about the relation *is greater than*, why don't syllogisms in Barbara depend for their validity on a comparable assumption about the relation *belongs to all*? As Frede 1974 notes,

...it is difficult to think of any satisfactory argument which would have shown that 'belonging to' is in a privileged position and at the same time would not have indicated that other expressions are in the same privileged position and which therefore would have forced the Peripatetics to admit arguments as syllogisms which they did not want to count as such. (108)

Such reflections would naturally lead to precisely the kind of argument Alexander attributes to the *neoterói*:<sup>27</sup>

1. Categorical syllogisms are relevantly similar to (GT) and (EQ).
2. (GT) and (EQ) depend for their validity on implicit axioms.
3. Therefore, categorical syllogisms depend for their validity on implicit axioms.
4. Hence, categorical syllogisms are not fully explicit.
5. Hence, they are not really syllogisms.

### A.2.3 The Peripatetic response

Alexander parries this argument with a block and a counterpunch. First he argues that categorical syllogisms are *not* relevantly similar to (EQ) and other unmethodically conclusive arguments. Then he argues that in fact it is Stoic *hypothetical* syllogisms that depend, like (EQ), on unstated assumptions.

What categorical syllogisms have that unmethodically conclusive arguments do not, Alexander claims, is universal premises. Without at least one universal premise, there can be no syllogism (345.18-20). The validity of unmethodically conclusive arguments with

---

<sup>27</sup>Cf. Barnes 1990:76.

particular premises rests on the truth of unexpressed universal premises (345.24-7). It is obvious, however, that this response begs the question. It is true that *categorical* syllogisms require a universal premise, but whether all syllogisms are categorical is precisely what is at issue. Alexander tries to free his argument from this assumption by arguing that no arguments from two particular premises could hold “in all matter” (*epi pasēs hulēs*, 345.20-2) and showing how to construct formal counterexamples to arguments like (EQ) (344.31-345.12, 348.5-12). But what counts as the *matter* in an argument is just as much at issue as what counts as a syllogism. If my conjecture is correct, the *neoteroi* are suggesting that “belongs to all” and “belongs to some” ought to be counted as matter, just as “is equal to” is. And if these are counted as matter—that is, as subject to replacement for the purposes of finding formal counterexamples—then no categorical syllogisms will hold “in all matter.”

Alexander’s counterpunch is more interesting. He attempts to show that (what we would regard as) valid arguments of propositional logic depend for their validity on unstated universal premises. The first he draws from Plato’s *Republic*:

- (PR) If he was the son of a god, he was not greedy.  
 If he was greedy, he was not the son of a god.  
 Therefore, he was not both [greedy and the son of a god].(22.8-9)

The Stoics (and presumably the *neoteroi*) would have counted this argument as syllogistically conclusive. But Alexander says that “it does not conclude through the premises laid down, but by the addition of a universal premise—namely, ‘when from each of a pair of contradictories follows the contradictory of the other, it is impossible for both to belong to the same thing’” (22.9-12). In assimilating this argument to “the arguments which the *neoteroi* call unmethodically conclusive” (22.18),<sup>28</sup> Alexander is suggesting a criticism of Stoic logic that is the exact mirror image of the *neoteroi*’s criticism of categorical syllogistic: Stoic “syllogisms” are like unmethodically conclusive arguments in that they depend

<sup>28</sup>I do not think that Alexander’s wording implies that the *neoteroi* themselves call (PR) “unmethodically conclusive.” He says only that the arguments the *neoteroi* call unmethodically conclusive are “of this sort.”

for their validity on an unstated assumption; hence they are not really syllogisms at all.

Similarly, Alexander suggests that the argument

- (HS) If it is human, it is animal.  
 If it is animal, it is substance.  
 Therefore, if it is human, it is substance.

is not syllogistic (347.18-20), but depends on the true universal assumption “everything that follows from something also follows from that from which the first thing follows” (347.26; Barnes 1990:114).<sup>29</sup> This is not a direct criticism of Stoic logic, since apparently the Stoics would not have counted (HS) as a syllogism (Frede 1974:106). But Alexander’s demand that we state what we would regard as laws of propositional logic as categorical propositions indirectly subverts Stoic logic, by suggesting that the inferences it regards as completely explicit in fact depend on unstated assumptions.

#### A.2.4 Galen’s pragmatic alternative

Both parties to the debate, as I have reconstructed it, argue that their opponents’ paradigm syllogisms really have the same status as the unmethodically conclusive arguments: they are conclusive, but only in virtue of some true assumption that they fail to make explicit. One response to this situation—a response I think we can see in Alexander’s contemporary Galen—is to conclude that *all* arguments depend on unstated assumptions. Galen belittles the intolerance of the Stoic and Peripatetic schools; he accepts both categorical and hypothetical syllogisms, as well as arguments like (EQ), which he places in a third category of “relational syllogisms” (*sullogismoι kata to pros ti*, IL xvi).<sup>30</sup> He criticizes the Peripatetics for trying to force relational syllogisms into categorical form (xvi.1). Though he acknowledges the possibility of putting relational syllogisms into either categorical or hypothetical form by the addition of a self-evident axiom (cf. xvi.11, xvi.5), he sees no point in doing

<sup>29</sup>Alexander notes at 373.29-35 that “A follows from B” is equivalent in his usage to “if B then A.” He criticizes the Stoics for distinguishing between what he regards as two ways of saying the same thing and claiming that a syllogism comes about only if the latter form is used. Incidentally, the conditionals in (HS) are expressed with genitive absolutes, not the word “if.”

<sup>30</sup>All translations from Galen are from Kieffer 1964.

so. On Galen's view, (EQ) and other relational syllogisms are perfectly good syllogisms as they stand.<sup>31</sup> True, "they have the cause of their structure (*sustaseōs*) derived from certain axioms" (xvi.5):<sup>32</sup> but this does not seem to distinguish them from other syllogisms, since "[n]early all the syllogisms get their structure through the cogency of the universal axioms that are set over them" (xvii.1).

Galen seems to be expressing the view that it is futile to ask when *all* of the assumptions on which an inference depends have been made explicit, and proposing instead that an inference has sufficiently explicated when its structure is determined by a self-evident axiom.<sup>33</sup> We *could* privilege hypothetical or categorical syllogisms, but there is no reason to do so; and once we allow them both, we might as well also count as syllogisms a whole slew of inferences that are neither categorically nor hypothetically valid as they stand.<sup>34</sup> Ebbesen sums up Galen's contribution as "the idea that all syllogizing depends on self-evident axioms, so that all syllogistic validity can be given a material explanation..." (1981:116). But perhaps it would be more correct to say that for Galen, the line between formal and

<sup>31</sup>In xvi.10 Galen calls the axiom a "conjoined axiom" (*sunēmmenou axiōmatos*), which suggests an explicitly added assumption (unless *sunēmmenon axiōma* at xvi.10 means "conditional proposition," as in the Stoic usage, v.5). But his examples of relational syllogisms (except xvii.3 and xvi.11) do not reflect this, and his claim (xvi.12) that the axiom accounts for "the credibility of [relational syllogisms'] *structure* and their demonstrative force" suggests otherwise (cf. i.3). Kieffer 1964 proposes, reasonably, that the axiom functions as a kind of argument schema (118), which *can* be made explicit, but which Galen thinks there is no *reason*, beyond pedantry, to make explicit (120). Once one accepts the idea that *all* forms of argument depend on axioms, the issue of which axioms to make explicit becomes a pragmatic one.

<sup>32</sup>Galen attributes the view that relational syllogisms are "conclusive by force of axiom" (*sunaktikous kata dunamin axiōmatos*) to Posidonius. How much of Galen's view should be attributed to Posidonius has been much debated (Mueller 1974:62, Ebbesen 1981:113, Barnes 1990:99 n. 207, Kidd 1978, Kieffer 1964:28-30).

<sup>33</sup>In i.5, Galen distinguishes *axiom* ("a proposition carrying conviction of itself to the intellect") from *premise* (a "statement about the nature of things," not self-evident, but derived from perception or demonstration). I do not think that Galen means to suggest that axioms are not about "the nature of things"; the distinction is rather an epistemological one (cf. xvii.7).

<sup>34</sup>The text of the last part of the *Institutio Logica* is corrupt, and interpretation is difficult. In particular, Galen's claim at xvi.5 that by keeping in mind the axioms on which relational syllogisms depend, "we shall be able to begin again more clearly and reduce such syllogisms to the categorical form" threatens to efface the difference between his view and Alexander's (Barnes 1990:99 n. 211). In xvi.11, Galen shows how a relational syllogism might be put into both hypothetical and categorical forms. (He shows no greater awareness than Alexander of the logical problems with the latter approach, taking "The man whom someone has as father, of him he is the son" as a universal categorical proposition.)

material is purely pragmatic: one should seek forms of argument that do not *need* further explication (because the axioms on which they depend are self-evident and will not be disputed), not forms of argument that *cannot* be further explicated. For *every* form of argument can be further explicated,<sup>35</sup> if one is willing to descend into needless pedantry.<sup>36</sup>

### A.3 The transformation of the topics

Of course, Abelard would not have had access to Galen, the Stoics, or the Greek commentators on Aristotle. He may have known Aristotle's *Prior Analytics*, but probably not the *Topics* (De Rijk 1956, xvi ff., Ebbesen 1982:104-9). His main sources for the logic of antiquity were Porphyry and especially Boethius, to whom he refers constantly in the *Dialectica*. But the debate over unmethodically conclusive arguments influenced him indirectly. Galen's view that relational syllogisms are valid in virtue of an unexpressed axiom, which need not be made explicit as part of the argument itself, probably influenced Themistius' "combination of axiomatics and topics" (Ebbesen 1981:117). In turn, Themistius' conception of the Aristotelian *Topics* was transmitted to the middle ages through Boethius and framed the debates about the source of the validity of the syllogism up to the time of Abelard.

#### A.3.1 Aristotle's Topics

Aristotle's *Topics* is a collection of general rules or heuristics for the construction of convincing and "reputable" (*endoxoi*) dialectical arguments on any subject whatever. It is notable that Topical arguments do not have the form of categorical syllogisms: their reputability cannot be read off from their syntactic structure, but depends on relations between terms.

---

<sup>35</sup>This claim would be better supported if we could be sure that Galen's claim at xvii.2 that *all* demonstrative syllogisms depend on axioms is meant to apply to categorical and hypothetical syllogisms. "Unfortunately," as Kieffer notes, "Galen does not expressly say this, and there are no examples given in this appendix of plain categorical or hypothetical syllogisms illustrating the rule" (123).

<sup>36</sup>Galen may have been influenced by the Sceptics' charge that the major (conditional) premise in a syllogism is redundant (Sextus Empiricus *PH* II.159-166). We know that Galen thought that much traditional logic was useless and even considered becoming a Sceptic (Ebbesen 1981:114-5; Kieffer 1964:1).

It is tempting to think that their reputability depends on Topical “maxims” which function like quasi-logical laws,<sup>37</sup> but this is almost certainly not Aristotle’s view. Aristotle does not seem to think that the acceptability of the particular syllogisms *depends on* the general rules he gives; if he did, it would be difficult to make sense of the fact that he often acknowledges counterexamples or *enstaseis* to the maxims (e.g., at 115b14, 117a18, 117b14, 121b30, 123b17, 124b19, 128b6).<sup>38</sup> More likely, the reputability of each particular syllogisms is to be judged case by case.<sup>39</sup> Thus it seems best to view the Topics as a collection of practical suggestions for finding reputable arguments in real dialectical disputes, rather than as norms for “material validity” (cf. Allen 1995:189, Stump 1978:168-177, Green-Pedersen 1984:23).

### A.3.2 Topics as axioms

Beginning with Themistius—and probably even earlier<sup>40</sup>—the Topical maxims began to be conceived as *axioms* on which the validity of dialectical arguments might rest. Themistius’ account of the Topics has survived only in a report by Averroes, but it is evident from that report that Boethius follows Themistius closely in his works on the Topics (*DTD*, *ITC*).<sup>41</sup> From Averroes and Boethius, we can reconstruct Themistius’ view as follows:

An Aristotelian *topos* is an axiom, that is, a self-evident, primitive, universal proposition. All arguments derive their force from such axioms. In some arguments the axiom is explicitly stated, in others it is implicit. (Ebbesen 1981:118)

---

<sup>37</sup>For example, the topical argument “since perceiving is judging, and it is possible to judge correctly and incorrectly, there would also be correctness and error in perception” (*Top* II.4.111a16-18) seems to depend on the maxim “if contrary attributes belong to the genus, they also belong to the species.” Many commentators have claimed that the Topics contain quasi-logical laws (see Brunschwig 1967:xl-xli, De Pater 1968:166, 174).

<sup>38</sup>In fact, at *Top* 155b29-35, Aristotle suggests that in some cases one will have to use instances to establish the general “law” (e.g., that knowledge of contraries is the same) by induction.

<sup>39</sup>This does not rule out the possibility that a syllogism may inherit some reputability or justification from the *Topos* that generates it, as Alexander seems to have thought (*Top* 126.23-4).

<sup>40</sup>Ebbesen 1981 conjectures that Themistius takes his conception of the Topics from an “unknown third century commentator” (117).

<sup>41</sup>The parallel between Averroes’ report and passages in Boethius’ *DTD* and *ITC* was first noticed by Stump (Ebbesen 1981:118, where the Averroes passage is quoted).

This is certainly a departure from Aristotle, who never describes his *Topics* as “axioms”—propositions that one must grasp in order to learn anything at all (*AnPo* I.2.72a16-17)—nor as self-evident or primitive, let alone universal (recall his bland acceptance of counterexamples).<sup>42</sup> Ebbesen conjectures that

... the strange initial identification of topical propositions with demonstrative axioms and the talk about the power (*vis* in Boethius) they lend to the arguments is best explained on the hypothesis that the Themistian theory of topical arguments arose when he or a predecessor saw that Galenic axiomatic proof had many similarities with Aristotelian topical proof. (1981:120)

Such an assimilation of Galenic axioms to Aristotelian Topical maxims would have been quite natural. A commentator persuaded by Galen to give up belief in the primacy of the categorical syllogism might reasonably look to the *Topics* for a source of general principles that might provide the basis for various forms of argument. The word “axiom” would have suggested itself in view of Posidonius’ claim that relational syllogisms are “conclusive by force of axiom” (Galen *IL* xviii.8). Moreover, Galen comments that the syllogism

The good of the better is worthier of choice.  
Soul is better than body.  
Therefore, [the good] of the soul is worthier of choice than that of the body,

which is nearly the same as one of Aristotle’s Topical syllogisms (*Top* III.2.117b33-9), is similar to his own relational syllogisms, which depend on axioms (*IL* xvi.13, Ebbesen 1981:117).

According to Boethius, the word “Topic” (Latin *locus*) can designate either a “maximal proposition” (*maxima propositio* or simply “maxim”) or a “Differentia” (1185A-1186B). “Maximal proposition” is Boethius’ translation of Themistius’ “axiom” (Ebbesen 1981:120): “... a maximal, universal, principal, indemonstrable, and known per se proposition, which in argumentations gives force to arguments and to propositions...” (*DTD* 1185B).<sup>43</sup> It is striking that Boethius’ first example of a maximal proposition is “if you take equals from

<sup>42</sup>Aristotle associates *axiōmata* with demonstration, not dialectic: see *AnPo* I.7.75a39-42, with Ross 1949’s note.

<sup>43</sup>All translations from *DTD* are from Stump 1978. Boethius notes that maxims need not be necessary: some are merely probable (*ITC* 1052B, Green-Pedersen 1984:62).

equals, the remainders are equal” (*DTD* 1176C)—precisely the kind of proposition that Galen would have regarded as an *axiom* grounding a relational syllogism.<sup>44</sup> Like Galen’s axioms, Topical maxims can either be contained in arguments as explicit premises or serve as external “guarantors of validity or of soundness” (Stump 1989:39; cf. Green-Pedersen 1984:69, *DTD* 1185B, *ITC* 1051). The Differentiae, on the other hand, are classes into which the maxims are grouped, depending on their terms: for example, “from substance,” “from opposites,” “from the whole,” “from similars” (1186A ff.). They function primarily as a tool for the *discovery* of maxims and intermediate terms appropriate to an argument (Stump 1978:195, 202; Green-Pedersen 1984:67).

For example (*DTD* 1188B-C), suppose we want to show that justice is advantageous. We might look to the Differentia “from the whole, that is, from genus,” noticing that the genus of justice is virtue. Using one of the maxims under this Differentia—“whatever is present to the genus is present to the species”—we can construct the following argument for the desired conclusion:

- (JA) Every virtue is advantageous.  
 Justice is a virtue.  
 Therefore, justice is advantageous.

In this example, the maxim is not included in the argument as a premise; we must therefore infer that it “supplies force to the argument and makes [it] complete from without” (1185B).

Commentators have wondered how the maxim here can play the role Boethius assigns it, of supplying force to the argument and completing it from without, when (JA) has the form of a valid categorical syllogism (Green-Pedersen 1984:68-9; cf. Stump 1978:1834, Abelard *D* 257.34-258.9).<sup>45</sup> But the problem only arises if we assume that valid categorical syllogisms are distinguished from other forms of arguments by the fact that they require no external

<sup>44</sup>It is also one of Aristotle’s examples of a “common” axiom: *AnPo* I.10.76a41.

<sup>45</sup>The example is not unique in this respect, but it is not typical either. Many of Boethius’ examples are neither categorical nor hypothetical syllogisms: for example, “if someone argues that the Moors do not have weapons, he will say they do not use weapons because they lack iron” (1189D). This appears to be a one-premise inference based on the maxim “where the matter is lacking, what is made from the matter is also lacking.”

validation.<sup>46</sup> And as we have seen, that assumption—dear to Alexander—had already been rejected by Galen. If Ebbesen is right that Boethian Topical inference is a descendent of Galen’s axiomatic proof, then Boethius’ use of (JA) as an example of an argument validated by a topical maxim is not so surprising. For Boethius as for Galen, Ebbesen claims, “*every* inference owes its cogency to an axiom” (emphasis added):

The implication of the Boethian theory would seem to be that all proof proceeds, implicitly or explicitly, by instantiation and detachment and, as some medievals saw, that a categorical syllogism is not anything *sui generis*, as it depends on a law of inference of the same type as the ones that licence inferences involving other relations than plain predication. (Ebbesen 1982:112)

It is not clear that Boethius himself accepts all these “implications.” But one can see how eleventh and twelfth century logicians—whose main sources for syllogistic theory were the works of Boethius—might have been led to them.

### A.3.3 Early medieval theories of Topics

What is implicit in Boethius becomes fully explicit in the earliest medieval theories of Topics. In Garlandus Compotista’s *Dialectica*, the theory of Topics is taken to be *prior* to the theory of categorical and hypothetical syllogistic: syllogisms are ratified by topical maxims (*per maximam propositionem syllogismus approbatur*, *D* 86.13).<sup>47</sup> For instance, the syllogism

---

<sup>46</sup>Stump 1978, who does make this assumption (183), proposes an interesting solution to the resulting problem. She notes that all of the syllogisms Boethius says are validated by an external maxim—and dialectical arguments generally—have *indefinite* premises: i.e., premises without explicit determinations of quantity (all/some). Hence, Stump claims, the syllogisms do not in fact have valid categorical forms (184-5), and that explains the need for an external ground. But it is unclear why the indefiniteness of the *minor* premise in (JA) should matter. Traditionally, indefinite premises are counted as particulars (Aristotle *AnPr* I.4.26a28-30 and Alex. *AnPr* 51.24-30—both of whom explicitly allow arguments of the same form as (JA) as valid first-figure categorical syllogisms). Moreover, Boethius says that “[a]rguments drawn from definition, genus, differentia, or causes most of all provide force and order to *demonstrative* syllogisms” (1195A-B, emphasis added): presumably demonstrative syllogisms will *not* have the indefinite premises characteristic of dialectical arguments, yet Boethius still seems to think that the maxims *from genus* will give them “force and order.” Finally, Galen (*IL* xiii.1) notes that the mood Barbara, which is “most appropriate to scientific demonstrations,” can be expressed in two forms: one with indefinite premises and one in which the quantity is explicitly marked.

<sup>47</sup>Green-Pedersen 1984 shows that Garlandus was anticipated in this view by Abbo of Fleury (945-1004) and other early commentators on the Boethian Topics (144, 152). He summarizes the

Every animal is a substance.  
 Every man is an animal.  
 Therefore, every man is a substance.

depends on the maximal proposition “that which is universally attributed to the whole is [also universally attributed] to the part” (*quod universaliter attribuitur toti, et parti*, 92.29). More generally, “categorical syllogisms are aided by the Topics *from the whole* and *from the part* and *from an equal*” (114.18). Topics apply to hypothetical syllogisms in two ways: both by providing the conditional major premise and by certifying the transition from the major and minor premise to the conclusion (Stump 1989:85). The validity of hypothetical syllogisms is secured by the maxims “when the antecedent is affirmed, the consequent is affirmed” (*posito antecedenti ponitur consequens*, 114.11) and “when the consequent is denied, the antecedent is denied” (*destructo consequenti destruitur antecedens*, 114.16)—that is, *modus ponens* and *modus tollens*. Stump sums up Garlandus’s view as follows: “. . . all inferences, whether in categorical or hypothetical syllogisms, are dependent on the Topics for their validity; and in the case of hypothetical syllogisms, the acceptability of the syllogism [that is, its soundness] is also Topically dependent” (1989:87; cf. 1982:277).

Interestingly, Garlandus acknowledges valid hypothetical syllogisms that do not proceed by either *modus ponens* or *modus tollens*: for example,

If it is a man, it is capable of laughter.  
 But it is not a man.  
 Therefore, it is not capable of laughter. (129.29-33)

This looks like a classic case of “denying the antecedent.” But the syllogism is valid, in Garlandus’s view, because the terms “man” and “capable of laughter” are “equals” (i.e., coextensive),<sup>48</sup> and a Topical maxim licenses the intersubstitution of coextensive terms pre-1100 works by saying that they take the Topics to be an “. . . ‘underlying logic’ which shows or explains why the arguments are valid. . .” (160).

<sup>48</sup>Garlandus seems to be following Boethius here (*ITC* 1132-3). In her note on the passage from Boethius, Stump writes: “In the case of the inferences involving *man* and *risible thing* mentioned here, the inferences are valid not in virtue of their form but in virtue of the meaning of the terms. In raising these latter inferences in the context of inferences of the first mode [i.e., *modus ponens*], Boethius gives the impression that he has not clearly distinguished these two different ways of warranting inferences” (1988:226).

(107.29-30).<sup>49</sup> Garlandus marks the difference between such syllogisms and standard hypothetical syllogisms by saying that the former conclude “from the force of the terms” (*ex vi . . . terminorum*, 129.27), while the latter conclude through *modus ponens* and *modus tollens*. But as Stump points out, the distinction is obscured by the fact that in *both* the standard and the exceptional cases, “the validity of hypothetical syllogisms . . . is guaranteed by maxims” (1989:83). If *all* syllogisms derive their validity from Topical maxims, then it is unclear what *significance* the distinction between syllogisms that are good *ex vi terminorum* and those we would regard as “formally valid” can have.

The early twelfth century works on the Topics collected in De Rijk 1962-7 seem to follow Garlandus in taking all syllogisms to be validated by topical maxims. The *Introductiones dialectice Berolinenses*, for instance, takes all syllogisms in the mood Barbara to be licensed by the following topical maxim:

If something is predicated universally of something, then if something else is predicated universally of the predicate, that same thing is predicated universally of the subject. (Stump 1989:116)

Similarly, *Abbreviatio Montana* presents a topical rule governing each valid syllogistic mood “in the same way it presented rules for the inferences in the preceding sections, assuming apparently that syllogistic inferences are one more variety of Topical inference” (Stump 1989:125). Stump concludes, on the basis of her reading of these early manuscripts, that Abelard was the first medieval logician to insist that the ground for the validity of categorical syllogisms is fundamentally different in character from the ground of such inferences as “Socrates is a man; therefore, Socrates is an animal:” that syllogisms are valid in virtue of their *construction* and do not depend at all on broadly metaphysical relations among things (1989:128).<sup>50</sup> After Abelard, on the other hand, the distinction is ubiquitous, though a few

<sup>49</sup>“de omnibus similibus, sive proportio sit sive paritas sive simplex similitudo, de similibus idem iudicium.”

<sup>50</sup>Besides the absence of this view in earlier writers, there is some direct testimony that Abelard (if he is to be identified with “Master P”) was its originator (Stump 1989:127, Green-Pedersen 1984:199-200). Stump suggests that Abelard may not have arrived at his mature view until his final revision of the *Dialectica* (in 1132-6) (130). Green-Pedersen 1984 makes the more cautious claim that “Abelard is the earliest author to go into a detailed and comprehensive discussion of the

commentaries still speak of applying Topics to syllogisms and a few actually argue against Abelard's distinction (Green-Pedersen 1984:198200).

Stump sums up the history nicely:

For Boethius dialectic was largely a corollary of metaphysics. The world has a certain nature, in consequence of which certain things are invariably or at least regularly connected with each other. Because we can know this nature and the variable or regular connections it involves, we can know that certain inferences among propositions preserve truth. In early medieval logic, before Abelard, no significant distinction was drawn between dialectic and the rest of logic in this regard; even logical laws warranting categorical syllogisms are treated as on a par with rules about the relationship between genera and species. Abelard tried to separate certain parts of logic from metaphysics by insisting that certain inferences hold not in virtue of any dialectical relationships but solely in virtue of their form. (1989:2)

Let us now turn to Abelard's view and arguments.

#### A.4 Abelard on perfect and imperfect inferentia

According to Abelard, there are two kinds of *inferentia*, or valid inference.<sup>51</sup> An inference is *perfect*, he says, when

...from the structure (*complexio*) of the antecedent itself, the truth of the consequent is manifest, and the construction (*constructio*) of the antecedent is so disposed that it contains also the construction of the consequent in itself, just as in syllogisms or in conditionals which have the form of syllogisms. (*D* 253.31-254.1)<sup>52</sup>

As an example of a perfect *inferentia* he offers a conditional formed from a categorical syllogism: "If every man is an animal and every animal is alive, every man is alive" (254.35). An inference is *imperfect*, by contrast, when the connection between antecedent and consequent takes its necessity "from the nature of things" (*ex rerum natura*, 255.7-8), not from problem" (194).

<sup>51</sup>Although Abelard is aware of the difference between arguments and conditionals, he applies the concepts *inferentia* and *consequentia*, as well as the perfect/imperfect distinction, to both (giving examples in both forms).

<sup>52</sup>"Perfecta quidem est inferentia, cum ex ipsius antecedentis complexione consequentis veritas manifesta est et antecedentis constructio ita est disposita, ut in se consequentis quoque constructionem contineat, veluti in syllogismis aut in his hypotheticis quae formas habent syllogismorum."

the construction of the antecedent and consequent, as in the inference: “If every man is an animal, every man is alive” (255.3). *Both* perfect and imperfect inferences require a necessary connection between antecedent and consequent—indeed, the sense of the consequent must be *contained in* the sense of the antecedent (283.37-284.8).<sup>53</sup> The difference is not in the strength of the modal connection (255.12-13), but in its ground.<sup>54</sup>

#### A.4.1 Topics and the grounds of inference

The function of a Topic, according to Abelard, is to confer inferential force on an *imperfect* consequence by grounding it in a real relation among the things to which its terms refer (256.35-257.1).<sup>55</sup> For example, the imperfect consequence “if it is a man, it is an animal” is justified by the Topic *from species*, since *man* is a species of *animal*, and we know that genus necessarily applies to species (257.4-5).<sup>56</sup> Following Boethius, Abelard takes a Topic to have two components: a *locus differentia* and a *maxima propositio*. The *locus differentia* (henceforth *Differentia*) is “that thing in the relation of which to something else the soundness of the entailment consists” (*ea res in cuius habitudine ad aliam firmitas consecutionis consistit*, 263.7-8).<sup>57</sup> In the example, the *Differentia* is *man* (*qua* species of animal). The

<sup>53</sup> “. . . non solum antecedens absque consequenti non potest esse verum, <sed etiam> ex se ipsum exigit. . .” See Kneale and Kneale 1962:217-8.

<sup>54</sup> Abelard claims that the reason for the clause “*per se ipsa*” (a translation of Aristotle’s “through their being so”) in the definition of the syllogism is to ensure that syllogisms are perfect (254). This connects his discussion with the hellenistic debate surveyed in section A.1 above.

<sup>55</sup> “Cuius quidem loci proprietates haec est: vim inferentiae ex habitudine quam habet ad terminorum illatum conferre consequentiae, ut ibi tantum ubi imperfecta est inferentia, locum valere confiteamur.” Note that Abelard does not think that the Topic *makes* the inference valid or “true.” For he holds that the truth of “if it is man, it is animal” does not depend on the existence of either man or animal: like all true consequences, it is an eternal truth (279.18). But if man and animal did not exist, then (as will be explained later in this paragraph) there would be no *locus differentia* and hence no Topical grounding. So the fact that man is species of animal cannot be the *cause* of the entailment (*consecutio*) but only its *proof* (*probatio*) (265.10-12). That is, the Topic is adduced *ad argumentum*, not *ad causam inferentiae*: the inference is true not *because* man is a species of animal, but *if* man is a species of animal (265.12-13). This suggests that what distinguishes perfect inferences from imperfect ones is a special *epistemic* character: their validity can be known independently of all knowledge about the world.

<sup>56</sup> “. . . ex ‘hominis’ habitudine ad ‘animal’—quia scilicet species eius est—valere constat inferentiam.”

<sup>57</sup> Although the *Differentiae* are things, not relations, they count as Topical *Differentiae* only insofar as they stand in relations to other things (Green-Pedersen 1984:167).

*maxima propositio* (henceforth *maxim*) is a general proposition justifying an inference from an antecedent proposition containing a term for the *Differentia* to a consequent proposition containing a term for the thing to which it is related. In the example, the *maxim* is “of whatever the species is predicated, so is the genus” (*de quocumque praedicatur species, et genus*, 263.18). The function of *Topics* is to ground imperfect inferences in real relations between things.

So far Abelard’s account of the Topical grounding of inferences accords with Boethius’. Abelard’s strikingly original move is to insist that *perfect* inferences do not stand in need of topical grounding at all.<sup>58</sup> His reason is that perfect consequences do not “take their truth...from the nature of things” (*veritatem...non ex rerum natura...tenent*, 256.21-2). We can see this independence from things, Abelard claims, by noting that perfect consequences remain true in “...whatever terms you substitute” (*qualescumque terminos apponas*, 255.32-3), whereas an imperfect consequence “depends on the nature of things” and does not “remain true in any terms whatsoever, but only in those which preserve the nature of the entailment” (356.8-10).<sup>59</sup> For example, the entailment in “if it is man, it is animal” can be destroyed by replacing “man” or “animal” with “stone” (356.15-19).

Therefore those consequences are correctly said to be true from the nature of things of which the truth varies together with the nature of things. But those [consequences] of which the construction preserves its necessity equally in any things at all, no matter what relations they have, take their truth from the construction (*complexione*), not from the nature of things ... (256.20-23)<sup>60</sup>

This is all that later medieval writers typically say about the distinction between formal and material consequence: formal consequences hold “in all terms.”<sup>61</sup> But Abelard cannot

<sup>58</sup> “...quia ita in se perfectae sunt huiusmodi inferentiae ut nulla habitudinis natura indigeant, nullam ex loco firmitatem habent” (256.34-5).

<sup>59</sup> “Ceterae quoque verae consequentiae, quarum inferentia ex rerum natura pendet, non in quorumlibet terminorum rebus verae consistunt, sed in his tantum quae naturam eius consecutionis servant.”

<sup>60</sup> “Istae ergo consequentiae recte ex natura rerum verae dicuntur quarum veritas una cum rerum natura variatur; illae vero veritatem ex complexione, non ex rerum natura, tenent quarum complexio necessitatem in quibuslibet rebus, cuiuscumque sint habitudinis, aequae custodit...”

<sup>61</sup> See Buridan TC:I.4.2-3, Pseudo-Scotus *In An. Pr.* I.10, 8ff., quoted in Bochenski 1961:30.12, Albert of Saxony *PL* IV.1, 24ra-b, quoted in Bochenski 1961:30.14.

stop here, for the dominant view at the time when Abelard is writing is that categorical syllogisms and other perfect inferences are grounded in Topics: Abelard even attributes this view to Boethius and Porphyry (257.32-258.13).<sup>62</sup> A proponent of such a view would not be fazed by the observation that syllogisms preserve validity in all substitution instances. Such a person might claim either that

1. for each substitution instance, there is a Topic grounding the inference in some relation in “the nature of things” or that
2. there is a single Topic that grounds *all* of the substitution instances in some very *general* features of things.

Abelard offers arguments against both approaches (258-262 in his treatment of inferences, 352-365 in his treatment of conditionals). It is a measure of the success of these arguments, I think, that they do not get repeated: it becomes customary in later medieval manuals to infer from an inference’s being good “in all terms” (i.e., schematic-formal) to its being good “in virtue of its construction” and not in virtue of the nature of things (3-formal).

#### A.4.2 Generality and abstraction from “the nature of things”

Abelard cannot take this inference for granted. In fact, he does not even think that the inference is sound. He claims that the inference “if it is alive, it is alive,” which certainly holds in all substitution instances, is *not* perfect in its construction (*ad inferentis constructionem*): one would have to add the premise “. . . and everything that is alive is alive” (255.19-27).<sup>63</sup> And he takes the inference from a conditional to its contrapositive to be dependant on the Topic “from an equal in inference” (351.29-352.11, Stump 1989:103). Evidently, then, there

<sup>62</sup>See section A.3.2, above.

<sup>63</sup>If Abelard thinks that two premises are always required for *perfectio constructionis*, as some passages suggest (e.g., 255.1-2), then he is going to have trouble with the *conversion* inferences necessary for the reduction of second- and third-figure syllogisms to the first figure. For these inferences—e.g. “all asses are animals, therefore some animals are asses”—have but a single premise. It would be awkward to maintain that the validation of second- and third-figure syllogisms, which are perfect in Abelard’s sense, requires the use of an imperfect inference. (So far as I am aware, Abelard does not discuss this problem.)

is more to perfection than mere preservation of validity “in all terms.” Abelard makes this point explicitly in his discussion of the hypothetical syllogisms Boethius takes to be valid by virtue of “the nature of the things, in which alone these propositions can be asserted” (*DHS* II.ii.4-5; see section A.1.3, above). After rejecting Boethius’ claim that inferences of the form “if it is not *a*, it is *b*; but it is *a*; thus it is not *b*” are valid in any terms for which the major premise can be asserted, Abelard goes on to say that *even if Boethius were right*, this fact would not show that such inferences are syllogisms (and hence perfect):

Even if it were possible, whenever the consequent were affirmed, necessarily to affirm the antecedent *from any property whatever*—nevertheless there would be no form of syllogism in which, the consequent having been affirmed in this way, one could affirm the antecedent, or the antecedent having been denied, one could deny the consequent, since the inference of a syllogism is supposed to be so perfect that no relation of things pertains to it. (502.19-25).<sup>64</sup>

To say that an inference is “perfect” is to say that our knowledge of its validity is completely independent of our knowledge of “the nature of things.” Even if Boethius were right that certain hypothetical inferences held in all terms for which the premises could be true, that would not be something we could know without knowing something about “the nature of things”—the relations in which various things stand. That an inference holds in all terms, then, is no guarantee that it abstracts entirely from the things those terms represent.

#### A.4.3 Abelard’s arguments for the 3-formality of syllogisms

Let us now consider Abelard’s arguments for the claim that syllogisms do not have Topical grounding. Recall that there are two ways in which one might oppose Abelard’s claim. First, one might argue that the validity of each individual syllogism is grounded in a particular relation between things (the “local strategy”). Second, one might argue that there is a single, very general relation between things that grounds the validity of *all* syllogisms in a particular mood (the “global strategy”). Abelard shows that neither approach will work.

<sup>64</sup>“Nec si etiam possit consequens positum necessario ponere antecedens ex quacumque proprietate, nulla tamen erit syllogismi forma, in qua hoc consequens positum ponat antecedens vel antecedens destructum destruat consequens, quippe syllogismi inferentia ita perfecta debet esse ut nulla rerum habitudo ad ipsam operetur...”

In my discussion, I will consider only categorical syllogisms, though Abelard brings similar considerations to bear on hypothetical ones.

### The local strategy

Given a particular categorical syllogism, the obvious place to look for a Topical Differentia is in the middle term. For example, in the syllogism

All animals are alive.  
All men are animals.  
Therefore, all men are alive.

one might naturally take *animal* to be the Differentia and apply the Topic *from the genus*, with the maxim “whatever is predicated of the genus is also predicated of the species.” But as Abelard points out, this Topic would only explain the inference from the *second* premise to the conclusion, not the inference from *both* premises together (258.14-17; cf. 356.4-11).<sup>65</sup>

Even this kind of Topical grounding will be impossible when syllogisms have *false* or *accidentally true* premises, for example:

Every body is colored.  
But everything sitting is a body.  
Therefore, everything sitting is colored (260.18-27).

In such a syllogism, “none of the propositions by themselves necessarily imply the conclusion” (*nulla propositionum ad conclusionem per se necessario antecedit*, 260.19-20). For there is no real relation in the nature of things that could license the transition from *either* of these premises to the conclusion.<sup>66</sup> *Body*, for instance, is not the genus of *sitting thing*, nor is *colored thing* the genus of *body*.<sup>67</sup> The only relation between terms to which we might appeal here is the relation of *predication*: *colored* is *universally predicated of body*, and *body*

<sup>65</sup> “. . . non quantum ad inferentiam totius syllogismi locum esse confitemur, non videlicet secundum hoc quod <ex> duabus simul antecedentibus propositionibus consequens infertur, sed quantum ad inferentiam unius antecedentium propositionum ad tertiam.”

<sup>66</sup> Similar considerations lead Abelard to claim that “if man is a species of stone, then if [something] is a man, it is a stone” is good in virtue of its construction (312). It could not take its necessity from “the nature of things,” because in the nature of things man is *not* a species of stone. (312-3)

<sup>67</sup> Abelard says at 285.20-29 that “if it is body, it is colored” is only accidentally true.

of *sitting thing* (cf. 259.1-9). But “A is universally predicated of B” might taken to express either

- (a) that A is *asserted of* all B (*secundum vocum enuntiationem*), or
- (b) that in the order of things, A is *true of* all B (*secundum rerum cohaerentiam*) (353.10-12).<sup>68</sup>

If it means merely (a) that A is *asserted of* all B, then it clearly cannot ground a necessary inference from “every C is B” to “every C is A”:

For who would concede that if “stone” were asserted universally of “man” in some assertion, whether true or false, the consequence which follows [i.e., ‘if every stone is an ass, then every man is an ass,’ 353.5] would be true? This is why we can assert “stone” (or anything else we like) of “man,” but our assertion, which is manifestly false, confers no truth on the consequence. (353.15-19)<sup>69</sup>

If, on the other hand, the relation “A is universally predicated of B” means that A is *true of* all B, then it is of no use in syllogisms with false premises, such as

All men are stones.  
All stones are asses.  
Therefore, all men are asses (353.5).

Nor is it of any use when it is merely *accidental* that A holds of all B, since valid inference must be necessary (cf. 362.30-1). There are some categorical syllogisms, then, for which no local topical maxim can be found. I trust that no one will be tempted to say that these categorical syllogisms hold in virtue of their construction, while others do not. For once we accept that *one* syllogism in Barbara holds in virtue of its construction, we might as well accept that *all* do (since all have the same construction).

<sup>68</sup>For this distinction, see 329.19-35.

<sup>69</sup>“Quis enim consequi concedat ut, si ‘lapis’ de ‘homine’ universaliter enuntietur quacumque enuntiatione, sive scilicet vera sive falsa, vera sit illa consequentia quae sequitur? Unde est quia ‘lapidem’ vel quidlibet de ‘homine’ possumus enuntiare. Sed nihil veritatis enuntiatio nostra consequenti hypotheticae confert, quae aperte falsa est.”

### The global strategy

If the validity of categorical syllogisms depends on a Topical maxim, then, it must be a maxim that captures the dependence of the conclusion on *both* the premises. Syllogisms in Barbara, for instance, seem to depend on the rule:

(CS) If B is predicated of A universally and C is predicated of B universally, then C is predicated also of A universally,<sup>70</sup>

where “predicated of” is taken *secundum rerum cohaerentiam*. Might (CS) be a Topical maxim that gives syllogisms in Barbara their inferential force?

Abelard’s strategy here is to argue that (CS), while perhaps a true *rule* (*regula*), is not a Topical *maxim*, because it lacks a corresponding Differentia (261.34-5, 265.25-266.2, Stump 1989:96, Green-Pedersen 1984:197). The argument that (CS) lacks a Differentia is basically the same as the argument (rehearsed above) that particular syllogisms lack a Differentia. The Differentia would have to be some thing (*res*) that stands in the relation *predicated universally of* to some term in the conclusion. The only obvious candidate is the middle term (B). But the fact that B is predicated of all A could at best explain the validity of the inference from *one* premise of the syllogism to the conclusion (from “every B is C” to “every A is C”), not the validity of the inference from *both* premises to the conclusion. And it explains this only if B is predicated of all A *truly* and *necessarily*: that is, only if A and B stand in some beefier relation than mere predication—say, genus and species (362.26-31). But we need to be able to find a Differentia in arguments with *false* premises, too.

Why should it matter whether or not (CS) has a corresponding Differentia and is thus a true *maxim*? Here Abelard is not as explicit as he might have been, but I think we can reconstruct his reasoning. He is trying to show that syllogisms are grounded in their construction alone, not in “the nature of things.” Apparently, he takes the fact that syllo-

<sup>70</sup>I have used schematic letters to make the principle clearer. Abelard uses pronouns: “si aliquid praedicatur de alio universaliter et aliud praedicatur de praedicato universaliter, illud idem praedicatur et de subiecto universaliter” (261.14-16). There is a corresponding principle for hypothetical syllogisms: “si aliquid infert aliud et id quod inferat existat, id quoque quod infertur necesse est existere” (261.25-6).

gisms do not depend on any genuine *maxims* to be sufficient grounds for this claim. Thus, although he does not deny that (CS) is true if and only if the syllogism

All A are B.  
 All B are C.  
 Therefore, all A are C.

is valid,<sup>71</sup> he denies that this equivalence shows that our knowledge of the syllogism's validity depends on how things are in the world. In order to understand Abelard's reasoning here, we need to understand why he thinks that only a genuine Topical maxim—one with a Differentia—can ground the validity of an inference in “the nature of things.”

I propose that Abelard is thinking along the following lines. A Topical maxim gives a rule for inference that is *based on its locus differentia*: that is, on some *thing (res)* in the world.<sup>72</sup> The inferential force (*vim inferentiae*) which a maxim brings to an imperfect inference comes from the relation in which the Differentia stands to some term in the conclusion of the inference (*ex habitudine quam habet ad terminum illatum*, 256.36-7). For example, in the valid consequence “if it is man, it is animal,” the inferential force comes from the relation (*species*) in which the Differentia (*man*) stands to *animal*. The Differentia, then,

<sup>71</sup>In this respect, (CS) fares better than an alternative *regula*, (CS\*): “If B is predicated of A universally, then if C is predicated of B universally, then C is predicated also of A universally” (*si aliquid praedicatur de aliquo universaliter, tunc si aliud praedicatur de praedicato universaliter, et de subiecto*, 352.31-3). (CS) and (CS\*) are not equivalent, because the law of exportation fails in Abelard's logic. In fact, Abelard argues, (CS\*) and the corresponding *regulae* for other syllogistic moods have many false instances (358.34-362.17). For in order for a conditional to be true, on Abelard's view, the sense of the antecedent must contain within itself the sense of the consequent (253.28-9, 284.1-2, Martin 1986, Stump 1989:105). A conditional like “if every man is a body, then every man is colored” fails to satisfy this condition, since the connection between the antecedent and the consequent depends on an accidental truth (that every body is colored), and not merely on their senses. Hence the conditional “if every body is colored, then if every man is a body, every man is colored” must also be false, since it has a true antecedent and a false consequent (361.25, 28-29). But this conditional is just an instance of (CS\*). Since there is no comparable argument against the truth of (CS), I focus on it instead.

<sup>72</sup>Green-Pedersen 1984:167. In the consequence “if it is man, it is animal,” the *locus differentia* is *man*; when Abelard calls the Topic “from species,” giving the *relation* in which the Differentia stands to something else, he is saying “from where the locus comes” (*unde sit locus*, 264.5-34). Green-Pedersen conjectures, plausibly, that Abelard insists that the Differentia be a *thing* and not the relation itself because the latter approach would make the relations (e.g., genus, species) into “independent realit[ies]” and contradict his nominalism (168).

is the thing (*res*) in the nature of which the validity of imperfect inferences is grounded.<sup>73</sup> A *regula* without a *Differentia*, then, although it might still be thought to ground the validity of inferences, could not ground it in “the nature of things,” as a maxim does.

To modern eyes, this reasoning appears to make an unwarranted assumption: that the totality of facts about “the nature of things” is exhausted by facts of the form

A is F, or  
A stands in the relation R to B.

Given this assumption, it follows from (CS)’s lack of a *Differentia* that (CS) is not a fact about “the nature of things” and must therefore depend for its truth on something else: the *construction* or *form* of the syllogism, the way it is put together in thought and language. But if we relax the assumption and count as facts about “the nature of things” facts with more logical complexity—

A, B, and C stand in the relation Q, or  
not both: { (all A are B and all B are C) and not (all A are C) }, or even  
for all A, B, and C: A, B, and C stand in the relation Q,

—then there is no longer any reason to think that (CS) is not a fact about “the nature of things,” and consequently no reason to think that syllogisms in Barbara do not depend on facts about the world: more general facts, to be sure, than most Topically grounded inferences, but no less facts about “the nature of things.” Granted, the entailment in a categorical syllogism cannot depend on the real relation of *one* thing to another; but might it not depend on some more complex feature of the world?

This question would become acute for Kant—for whom “the nature of things” consists of just the kind of complex, generalized relational facts Abelard does not consider (e.g., the

<sup>73</sup>Cf. 255.7-9, on the consequence “if every man is animal, every man is alive”: “These inferences, although they are imperfect in the construction of the antecedent, nonetheless most often take their necessity *from the nature of things*, just as with [the consequence] which we put down earlier from ‘animal’ to ‘alive,’ since *the nature of animal*, in which as a substantial form *alive* inheres, never allows animal itself to exist without life.” (“Quae quidem inferentiae, quamvis imperfectae sint quantum ad antecedentis constructionem, tamen necessitatem ex rerum natura saepissime tenent veluti ista quam prius posuimus de ‘animali’ ad ‘animatum’, cum videlicet natura animalis, cui animatum ut substantialis forma inest, ipsum animal praeter animationem existere nusquam patiatur.”).

laws of Newtonian science)—and even more pressing for Frege, Russell, and Wittgenstein, whose new logical notation allowed the question to be raised in a more explicit way. But Abelard doesn't answer it. He is not even in a position to ask it. In order to do so, he would have to reject the broadly Aristotelian ontology he inherits from his sources and shares with all of his contemporaries. He would have to abandon the assumption that all facts about the world predicate “something of something” (*ti kata tinos*).<sup>74</sup> Given that assumption, Abelard is right to deny that syllogisms depend for their validity on facts about the world.

Indeed, the same reasoning that leads Abelard to this conclusion should lead him to accept the inference

- (EO)    A is east of B.  
           B is east of C.  
           Therefore A is east of C.

as valid in virtue of its construction.<sup>75</sup> For suppose the premises were false. What would be the Differentia? Since the inference is not valid in virtue of B's relation to something else, Abelard would reason, it must not be valid in virtue of “the nature of things.”

This point reveals the extent to which Abelard's arguments for the 3-formality of syllogisms are unavailable to us today. Abelard would have to concede that (EO) is valid in virtue of its construction, while

- (MA)    A is a man.  
           Therefore, A is an animal.

is valid in virtue of the nature of *man*. No modern advocate of logical hylomorphism, I take it, would make a principled distinction between these two cases. Similarly, as we have seen, Abelard takes syllogisms in Barbara to be valid in virtue of their construction, while denying the same status to

- (Id)    A is alive.  
           A is alive.

---

<sup>74</sup>On this, see Brower 1998: “According to Abelard, if a statement of the form ‘xRy’ is true, then what makes it true is nothing but individual subjects and their monadic properties” (623).

<sup>75</sup>I am not aware of any passages in which Abelard discusses such inferences.

and

(Cont) If A then B.  
If not B, then not A.

Again, his views about the basis for logical hylomorphism—views we do not share—lead him to make a distinction of principle where we see none.

To sum up: Abelard does not make the simple-minded argument that because valid syllogisms are valid in all substitution instances, their validity does not depend on the nature of things. That argument would not have been plausible in his historical context. Instead, he takes considerable pains to rebut versions of the view that syllogistic validity is Topically grounded. But in the end, his argument depends on tacit ontological premises about what can count as part of “the nature of things”—premises we no longer accept.

## A.5 Formal and material consequence

Abelard’s arguments seem to have been persuasive: according to Green-Pedersen 1984, the majority of Abelard’s twelfth century successors distinguish between “arguments which rest upon loci [Topics] (*locales*) and those that are valid by their form (*complexionales*)” (200). The distinction persists in the thirteenth century and is a likely ancestor of the fourteenth century (continental) distinction between formal and material consequence (Kneale and Kneale 1962:279), though the lines of influence are obscure. Here I can do little but sketch some of the later developments: more detail can be found in Kneale and Kneale 1962, Stump 1982, and Green-Pedersen 1984.

In their practice, thirteenth century writers follow Abelard in drawing a sharp distinction between syllogisms and most other inferences. Sometimes they even make the distinction explicit as a distinction between inferences valid in virtue of their construction and inferences valid in virtue of the terms or the nature of things (Green-Pedersen 1984:254). Yet after Abelard, no one seems to have cared much about the *basis* for the distinction. At any rate, thirteenth century commentators show little theoretical interest in the question of

which arguments are grounded in Topics (Green-Pedersen 1984:253). They typically see the Topics as instruments for the reduction of enthymemes to valid categorical syllogisms, and no longer as grounding the validity of inferences, but they do not explain why the categorical syllogisms do not *themselves* stand in need of further reduction.

In the fourteenth century, it becomes common to take *all* categorical syllogisms as dependent on the Topic “from a quantitative whole,” which is taken to justify the principle *dici de omni et nullo* (Stump 1982:287, 293; cf. Green-Pedersen 1984:256-7, 269). At the same time, Topical arguments begin to be referred to as “consequences” (*consequentiae*), and the role of consequences (e.g., conversion inferences) in syllogistic theory gives the theory of consequences a kind of priority over syllogistic (Stump 1982:290-3). In short, there is a reawakening of interest in non-syllogistic forms of inference and a blurring of the boundary between Topical arguments and syllogisms.

In this climate, two different distinctions come to be marked by the terminology of “formal and material consequence.” In England, logicians like Ockham and Burley distinguish formal consequences as those in which the antecedent is relevant to (or “contains”) the consequent, as opposed to consequences that are good in virtue of the material impossibility of the antecedent or the material necessity of the consequent.<sup>76</sup> A counterpart of Abelard’s distinction between perfect and imperfect *inferentia* is still used (in Ockham’s terminology, inferences valid *gratia formae* and *gratia materiae*), but it retains little of its former epistemic and metaphysical importance. On the continent (Pseudo-Scotus, Buridan, Albert of Saxony), the distinction between formal and material consequence closely resembles Abelard’s distinction between perfect and imperfect *inferentia*, but it is never given the kind of motivation that Abelard offers. The reason, perhaps, is that there is no longer a concerted opposition. After Abelard, it is taken for granted that valid inferences divide into those whose validity can be attributed to their structure and those whose validity depends on their terms and the nature of the things to which they refer.

I do not want to rest much on speculative claims about the influence of Abelard’s

<sup>76</sup>On the later development of this tradition, see Normore 1993:450-1, Ashworth 1982.

arguments on the later tradition. The important point is that Abelard's defense of the distinction between perfect and imperfect *inferentiae*—the deepest and fullest discussion we have of the basis for the widespread medieval distinction between arguments that are valid in virtue of their structure and those that depend on their terms and the nature of things—has its place in a philosophical framework we no longer share.

**Time Line**

300s	Aristotle, Theophrastus
200s	Chrysippus (Stoic logic)
100s	<i>neoterói?</i> (Mueller 1974:63)
0s B.C.	Posidonius, Boethus of Sidon, Cicero
0s A.D.	
100s	Galen, Alexander of Aphrodisias, Sextus Empiricus
200s	Plotinus, Porphyry
300s	Themistius
400s	Proclus
500s	Ammonius, Philoponus, Boethius
600s	
700s	
800s	
900s	Abbo of Fleury
1000s	Garlandus Compotista
1100s	<i>Introductiones Dialectice Berolinenses, Abbreviatio Montana,</i> Michael of Ephesus, Peter Abelard
1200s	Thirteenth-century commentators
1300s	Walter Burley, William Ockham, John Buridan, Albert of Saxony, Pseudo-Scotus

# BIBLIOGRAPHY

# BIBLIOGRAPHY

- Abelard, Peter *D* [1132–6]. *Dialectica*. See De Rijk 1956.
- Alberti, A., ed. 1990. *Logica, Mente e Persona*. Florence: Leo S. Olschki.
- Aldrich, Henry 1691. *Artis Logicae Compendium*. Oxford.
- Alexander of Aphrodisias 1991. *On Aristotle's Prior Analytics I.1–7*. Trans. Jonathan Barnes, Susanne Bobzien, Kevin Flannery, and Katerina Ierodiakonou. Ithaca: Cornell University Press.
- Alexander of Aphrodisias *AnPr*. In *Aristotelis Analyticorum Priorum Librum I Commentarium (Commentaria in Aristotelem Graeca II 1)*. Ed. M. Wallies. Berlin: G. Reimeri, 1883. Partial translation in Alexander 1991.
- Alexander of Aphrodisias *Top*. In *Aristotelis Topicorum Libros Octo Commentaria (Commentaria in Aristotelem Graeca II 2)*. Ed. M. Wallies. Berlin: G. Reimeri, 1881.
- Allgemeine Deutsche Biographie*. Leipzig: Verlag von Duncker and Humblot, 1876–1912.
- Allen, James 1995. The Development of Aristotle's Logic: Part of an Account in Outline. *Proceedings of the Boston Area Colloquium in Ancient Philosophy* 11, ed. John J. Cleary and William C. Wians, 177–205.
- Allison, Henry 1973. *The Kant-Eberhard Controversy*. Baltimore: Johns Hopkins University Press.

- Allison, Henry E. 1985. The Originality of Kant's Distinction between Analytic and Synthetic Judgements. In *Studies in Philosophy and the History of Philosophy*, vol. 12: *The Philosophy of Immanuel Kant*. Ed. Richard Kennington. Washington, D.C.: The Catholic University of America Press, 15–38. Reprinted in Chadwick and Cazeaux 1992, 324–345.
- Almog, Joseph, John Perry, and Howard Wettstein 1989. *Themes from Kaplan*. Oxford: Oxford University Press.
- Ammonius *AnPr*. In *Aristotelis Analyticorum Priorum Librum I Commentarium (Commentaria in Aristotelem Graeca IV 6)*. Ed. M. Wallies. Berlin: G. Reimeri, 1899.
- Anderson, A. R., N. D. Belnap, and J. M. Dunn 1992. *Entailment: the Logic of Relevance and Necessity*. Vol. 2. Princeton: Princeton University Press.
- Aristotle *AnPo*. *Posterior Analytics*. Greek text in Ross 1949.
- Aristotle *AnPr*. *Prior Analytics*. Greek text in Ross 1949.
- Aristotle *SE*. *Sophistical Refutations*. Greek text in Ross 1958.
- Aristotle *Top*. *Topics*. Greek text in Ross 1958.
- Arnauld, A. and P. Nicole 1662. *La Logique, ou l'art de penser* (known as the “Port-Royal Logic”). Paris. Translation: Arnauld and Nicole 1996.
- Arnauld, A. and P. Nicole 1996. *Logic or the Art of Thinking*. Ed. and trans. Jill Vance Buroker. Cambridge: Cambridge University Press.
- Ashworth, E. J. 1982. The eclipse of medieval logic. In Kretzmann, Kenny, and Pinborg 1982, 787–796.
- Aspray, W., and P. Kitcher 1988. *History and Philosophy of Modern Mathematics: Minnesota Studies in the Philosophy of Science* 11. Minneapolis: University of Minnesota Press.

- Ayer, Alfred Jules 1946. *Language, Truth, and Logic*. Second edition. London: Gollantz.
- Barnes, Jonathan 1990. Logical Form and Logical Matter. In Alberti 1990.
- Barwise, J. and S. Feferman, eds. 1985. *Model-Theoretic Logics*. New York: Springer-Verlag.
- Baumgarten, Alexander Gottlieb 1761. *Acroasis Logica in Christianum L. B. de Wolff*. In *Christian Wolff: Gesammelte Werke*, III.5, ed. J. École et al. Hildesheim and New York: Georg Olms, 1973.
- Beck, Lewis White 1955. Can Kant's Synthetic Judgments Be Made Analytic? *Kant-Studien* 47, 168–81.
- Belaval, Yvon 1960. *Leibniz Critique de Descartes*. Paris: Gallimard.
- Bell, David and Neil Cooper, eds. 1990. *The Analytic Tradition: Meaning, Thought and Knowledge*. Oxford: Blackwell.
- Beiser, Frederick C. 1987. *The Fate of Reason: German Philosophy from Kant to Fichte*. Cambridge, MA: Harvard University Press.
- Belnap, Nuel 1990. Declaratives Are Not Enough. *Philosophical Studies* 59, 1–30.
- Belnap, Nuel 1992. Branching Space-Time. *Synthese* 92, 385–434.
- Belnap, Nuel *UCL*. Under Carnap's Lamp: Value-meaning pre-semantics. Manuscript, 1999.
- Belnap, Nuel and Mitchell Green 1994. Indeterminism and the Thin Red Line. In *Philosophical Perspectives* 8, ed. James E. Tomberlin, 365–88.
- Belnap, Nuel and Gerald J. Massey 1990. Semantic Holism. *Studia Logica* 49, 67–82.
- Benacerraf, Paul 1973. Mathematical Truth. *Journal of Philosophy* 70, 661–80. Reprinted in Benacerraf and Putnam 1983, 403–420. (Citations are to reprint.)

- Benacerraf, Paul and Hilary Putnam, eds. 1983. *Philosophy of Mathematics: Selected Readings*. Second ed. Cambridge: Cambridge University Press.
- Beth, E. W. 1957. Le Savoir Dédectif dans la Pensée Cartésienne. In *Cahiers de Royaumont Philosophie*, II: *Descartes*. Paris: Minuit.
- Beth, Evert W. 1970. *Aspects of Modern Logic*. Dordrecht: D. Reidel.
- Bird, Alexander 1997. The Logic in Logicism. *Dialogue* 36, 341–60.
- Bochenski, I. M. 1956. *Formale Logik*. Freiburg and Munich: Karl Alber. Translation: Bochenski 1961.
- Bochenski, I. M. 1961. *A History of Formal Logic*. Trans. I. Thomas. Notre Dame: Notre Dame University Press.
- Boethius *DHS*. *De hypotheticis syllogismis*. See Obertello 1969.
- Boethius *DTD*. *De Topicis Diferentiis*. See Stump 1978.
- Boethius *ITC*. In *Topica Ciceronis Commentariorum Libri Sex*. See Stump 1988.
- Bolzano, Bernard 1837. *Wissenschaftslehre*. Cited from second ed., ed. Wolfgang Schultz. Leipzig: Felix Meiner, 1929. Partial translation: Bolzano 1972.
- Bolzano, Bernard 1972. *Theory of Science*. Ed. and trans. Rolf George. Berkeley: University of California Press.
- Boole, George 1854. *The Laws of Thought*. London: Macmillan. Reprinted New York: Dover, 1958.
- Boole, George 1868. Of Propositions Numerically Definite. *Transactions of the Cambridge Philosophical Society* 11, part II. Reprinted in Boole 1952.
- Boole, George 1952. *Studies in Logic and Probability*. La Salle, IL: Open Court.

- Boole, George 1997. *Selected Manuscripts on Logic and its Philosophy*. Ed. Ivor Grattan-Guinness and Gérard Bornet. Basel: Birkhäuser Verlag.
- Boolos, George 1975. On Second-Order Logic. *Journal of Philosophy* 72, 509–527. Reprinted in Boolos 1998, 37–53.
- Boolos, George 1984. To Be is to Be a Value of a Variable (or to Be Some Values of Some Variables). *Journal of Philosophy* 81, 430–449. Reprinted in Boolos 1998, 54–72.
- Boolos, George 1985. Reading the *Begriffsschrift*. *Mind* 94, 331–344. Reprinted in Boolos 1998, 155–170. (Citation is to reprint.)
- Boolos, George 1987. The consistency of Frege’s *Foundations of Arithmetic*. In J. J. Thomson, ed., *On Being and Saying: Essays for Richard Cartwright* (Cambridge, MA: MIT Press), 3–20. Reprinted in Boolos 1998, 183–201. (Citation is to reprint.)
- Boolos, George 1998. *Logic, Logic, and Logic*. Ed. Richard Jeffrey. Cambridge, MA: Harvard University Press.
- Bradley, F. H. 1883. *Principles of Logic*. London: Kegan Paul and Trench.
- Brandom, Robert 1994. *Making It Explicit: Reasoning, Representing, and Discursive Commitment*. Cambridge, MA: Harvard University Press.
- Brower, Jeffrey E. 1998. Abelard’s Theory of Relations: Reductionism and the Aristotelian Tradition. *Review of Metaphysics* 51, 605–631.
- Brunschwig, Jacques 1967, ed. and trans. *Aristote: Topiques*. Tome I. Paris: Société d’Édition “Les Belles Lettres.”
- Buickerood, James G. 1985. The Natural History of the Understanding: Locke and the Rise of Facultative Logic in the Eighteenth Century. *History and Philosophy of Logic* 6, 157–190.

- Buridan, John *TC* 1976. *Tractatus de Consequentibus*. Ed. Hubert Hubien. Louvain: Publications Universitaires.
- Burnyeat, Myles [A]. A Map of Metaphysics Zeta. Manuscript, 1995.
- Carnap, Rudolf 1934. Formalwissenschaft und Realwissenschaft. *Erkenntnis* 5. Translation: Formal and Factual Science, in Feigl and Brodbeck 1953.
- Carnap, Rudolf 1937. *The Logical Syntax of Language*. London: Routledge. Trans. Amethe Smeton from *Logische Syntax der Sprache* (Vienna, 1934). (Note that the 1934 edition was somewhat abridged.)
- Carnap, Rudolf 1942. *Introduction to Semantics*. Cambridge, MA: Harvard University Press.
- Carnap, Rudolf 1943. *Formalization of Logic*. Cambridge, MA: Harvard University Press.
- Carnap, Rudolf 1947. *Meaning and Necessity*. Chicago: University of Chicago Press.
- Chadwick, Ruth F. and Clive Cazeaux, eds. 1992. *Immanuel Kant: Critical Assessments*. Vol. II. London and New York: Routledge.
- Chihara, Charles 1998. Tarski's Thesis and the Ontology of Mathematics. In *The Philosophy of Mathematics Today*, ed. Matthias Schirn. Oxford: Clarendon Press.
- Church, Alonzo 1940. A Formulation of the Simple Theory of Types. *Journal of Symbolic Logic* 5, 56–68.
- Church, Alonzo 1944. Review of Carnap 1943. *Philosophical Review* 53, 493–8.
- Church, Alonzo 1956. *Introduction to Mathematical Logic*. (Revised ed.: first ed. 1944.) Princeton: Princeton University Press.
- Coffa, J. Alberto 1991. *The Semantic Tradition from Kant to Carnap to the Vienna Station*. Ed. Linda Wessels. Cambridge: Cambridge University Press.

- Conant, James 1991. The Search for Logically Alien Thought: Descartes, Kant, Frege, and the *Tractatus*. *Philosophical Topics* 20, 115–180.
- Corcoran, John 1974, ed. *Ancient Logic and Its Modern Interpretations*. Dordrecht: D. Reidel.
- Couturat, Louis 1901. *La Logique de Leibniz*. Paris. Reprinted Hildesheim: Georg Olms, 1961.
- Cresswell, M. J. 1973. *Logics and Languages*. London: Methuen and Co.
- Curry, Haskell B. 1951. *Outlines of a Formalist Philosophy of Mathematics*. Amsterdam: North Holland.
- Curry, Haskell B. 1963. *Foundations of Mathematical Logic*. New York: Dover, 1976. (Orig. ed. 1963).
- De Morgan, Augustus 1858. On the Syllogism: III. *Transactions of the Cambridge Philosophical Society* 10, 173–230. Reprinted in De Morgan 1966, 74–146. (Citation is to reprint.)
- De Morgan, Augustus 1860. Syllabus of a Proposed System of Logic. (Manuscript.) In De Morgan 1966.
- De Morgan, Augustus 1860b. Logic. From the *English Cyclopaedia* V. Cited from De Morgan 1966.
- De Morgan, Augustus 1966. *On the Syllogism (and Other Logical Writings)*. Ed. Peter Heath. New Haven: Yale University Press.
- De Pater, W. A. 1968. La fonction du lieu et de l'instrument dans les *Topiques*. In Owen 1968.
- De Rijk, L. M. 1956. *Petrus Abaelardus Dialectica*. Assen: Van Gorcum.
- De Rijk, L. M. 1959. *Garlandus Compotista Dialectica*. Assen: Van Gorcum.

- De Rijk, L. M. 1962–7. *Logica Modernorum*. Assen: Van Gorcum.
- de Vleeschauwer, Herman-J. 1939. *L'évolution de la pensée Kantienne*. Paris: Presses Universitaires de France. Citations are to the translation: *The Development of Kantian Thought: The History of a Doctrine*. Trans. A. R. C. Duncan. London: Thomas Nelson and Sons, 1962.
- Demopoulos, William 1994. Frege and the Rigorization of Analysis. *Journal of Philosophical Logic* 23, 225–46. Reprinted in Demopoulos ed. 1995, 68–88.
- Demopoulos, William, ed. 1995. *Frege's Philosophy of Mathematics*. Cambridge, MA: Harvard University Press.
- Descartes, René 1985. *The Philosophical Writings of Descartes*. Trans. John Cottingham, Robert Stoothoff, Dugald Murdoch. Volume 1. Cambridge: Cambridge University Press.
- Descartes, René *P* [1644]. *Principia Philosophiae*. Amsterdam: Elzevir. Translated as *Principles of Philosophy* in Descartes 1985.
- Descartes, René *R* [1628/1701]. *Regulae ad Directionem Ingenii*. Published posthumously, Amsterdam: P. and J. Blaeu. Translated as *Rules for the Direction of the Mind* in Descartes 1985.
- Diogenes Laertius *LP*. *Lives of Eminent Philosophers*. Two vols. Trans. R. D. Hicks. Cambridge, MA: Harvard University Press.
- Došen, Kosta 1994. Logical Constants as Punctuation Marks. In Gabbay 1994, 273–296.
- Dreben, Burton and Juliet Floyd 1991. Tautology: How Not to Use a Word. *Synthese* 87, 23–49.
- Dummett, Michael 1959. Truth. *Proceedings of the Aristotelian Society* n.s. 59, 141–62. Reprinted with Postscript in Dummett 1978, 1–24.

- Dummett, Michael 1973. *Frege: Philosophy of Language*. Cambridge, MA: Harvard University Press.
- Dummett, Michael 1978. *Truth and Other Enigmas*. Cambridge, MA: Harvard University Press.
- Dummett, Michael 1981. *The Interpretation of Frege's Philosophy*. Cambridge, MA: Harvard University Press.
- Dummett, Michael 1991. *The Logical Basis of Metaphysics*. Cambridge, MA: Harvard University Press.
- Ebbesen, Sten 1981. *Commentators and Commentaries on Aristotle's Sophistici Elenchi*. Volume I (The Greek Tradition). Leiden: E. J. Brill.
- Ebbesen, Sten 1982. Ancient Scholastic Logic as the Source of Medieval Scholastic Logic. In Kretzmann et al. 1982, 101–127.
- Einstein, Albert 1921. *Geometrie und Erfahrung*. Trans. G. B. Jeffery and W. Perrett, as Geometry and Experience, in Einstein 1983. (Citation is to translation.)
- Einstein, Albert 1983. *Sidelights on Relativity*. New York: Dover.
- Etchemendy, John 1983. The Doctrine of Logic as Form. *Linguistics and Philosophy* 6, 319–34.
- Etchemendy, John 1988a. Models, Semantics, and Logical Truth. *Linguistics and Philosophy* 11, 91–106.
- Etchemendy, John 1988b. Tarski on Truth and Logical Consequence. *Journal of Symbolic Logic* 53, 51–79.
- Etchemendy, John 1990. *The Concept of Logical Consequence*. Cambridge, MA: Harvard University Press.

- Euclid *E. The Thirteen Books of Euclid's Elements*. Trans. (with commentary) Thomas L. Heath. Volume One. New York: Dover, 1956.
- Evans, Gareth 1976. Semantic Structure and Logical Form. In Evans and McDowell 1976. Reprinted in Evans 1985, 49–75, with an Appendix, 405–7. (Citation is to reprint.)
- Evans, Gareth 1985. *Collected Papers*. Ed. John McDowell. Oxford: Clarendon Press.
- Evans, Gareth, and John H. McDowell 1976. *Truth and Meaning: Essays in Semantics*. Oxford: Clarendon Press, 1976.
- Feferman, Solomon [A]. Logic, Logics, and Logicism. Lecture, Boolos Symposium, University of Notre Dame, April 16–18, 1998. Manuscript, 2000.
- Feigl, H., and Brodbeck, M., eds. 1953. *Readings in the Philosophy of Science*. New York: Appleton-Century-Crofts.
- Field, Hartry 1984. Critical Notice of Crispin Wright: *Frege's Conception of Numbers as Objects*. *Canadian Journal of Philosophy* 14, 637–62. Reprinted in Field 1989 (Platonism for Cheap? Crispin Wright on Frege's Context Principle), 147–170.
- Field, Hartry 1989. *Realism, Mathematics and Modality*. Oxford: Basil Blackwell.
- Fine, Kit 1998. The Limits of Abstraction. In Schirn ed. 1998, 503–629.
- Flynn, Bernard M. 1946. The Notion of Formal Logic. *Laval Théologique et Philosophique* 2:1, 181–3.
- Føllesdal, Dagfinn 1980. Comments on Quine [1980]. In Kanger and Öhman 1980, 29–35.
- Frede, Michael 1974. Stoic vs. Aristotelian Syllogistic. *Archiv für Geschichte der Philosophie* 56. Reprinted in Frede 1987, 99–124. (Citation is to reprint.)
- Frede, Michael 1987. *Essays in Ancient Philosophy*. Minneapolis: University of Minnesota Press.

- Frege, Gottlob 1953. *The Foundations of Arithmetic: A logico-mathematical enquiry into the concept of number*. Trans. J. L. Austin. Oxford: Blackwell.
- Frege, Gottlob 1964. *The Basic Laws of Arithmetic: Exposition of the System*. Trans. M. Furth. Berkeley: University of California Press.
- Frege, Gottlob 1969. *Nachlass*. Ed. Hans Hermes, Friedrich Kambartel, Friedrich Kaulbach. Hamburg: Felix Meiner.
- Frege, Gottlob 1972. *Conceptual Notation and Related Articles*. Trans. and ed. Terrell Ward Bynum. Oxford: Clarendon Press.
- Frege, Gottlob 1984. *Collected Papers on Mathematics, Logic, and Philosophy*. Ed. Brian McGuinness. Oxford: Blackwell.
- Frege, Gottlob *ACN* [1882–3]. Über den Zweck der Begriffsschrift (On the Aim of the “Conceptual Notation”). *Sitzungsberichte der Jenaischen Gesellschaft für Medizin und Naturwissenschaft, Jenaische Zeitschrift für Naturwissenschaft* 16, 1–10. Translation by T. W. Bynum in Frege 1972. (Citation is to pagination in the translation.)
- Frege, Gottlob *BGS* [1879]. *Begriffsschrift, eine der arithmetischen nachgebildete Formelsprache des reinen Denkens* (Conceptual Notation: a formula language of pure thought modelled upon the formula language of arithmetic). Halle: L. Nebert, 1879. Translation by T. W. Bynum in Frege 1972.
- Frege, Gottlob *C*. *Philosophical and Mathematical Correspondence*. Ed. Gottfried Gabriel et al. Oxford: Basil Blackwell, 1980.
- Frege, Gottlob *CES* [1895]. Kritische Beleuchtung einiger Punkte in E. Schröders Vorlesungen über die Algebra der Logik (A Critical Elucidation of some Points in E. Schröder, *Lectures on the Algebra of Logic*). *Archiv für systematische Philosophie* 1, 433–56. Translation by Peter Geach in Frege 1984.

- Frege, Gottlob *CT* [1923]. Logische Untersuchungen; Dritter Teil: Gedankengefüge (Compound Thoughts). *Beiträge zur Philosophie des deutschen Idealismus* 3, 36–51. Translated by Peter Geach and R. H. Stoothoff in Frege 1984.
- Frege, Gottlob *FA* [1884]. *Die Grundlagen der Arithmetik: eine logisch mathematische Untersuchung über den Begriff der Zahl*. Breslau: Willhelm Koebner. Translation, Austin 1953.
- Frege, Gottlob *FC* [1891]. *Funktion und Begriff* (Function and Concept). Jena: H. Pohle. Translation by Peter Geach in Frege 1984.
- Frege, Gottlob *FG2* [1906]. Über die Grundlagen der Geometrie II (On the Foundations of Geometry: Second Series). *Jahresbericht der Deutschen Mathematiker-Vereinigung* 15, 293–309, 377–403, 423–30. Translation by E.-H. W. Kluge in Frege 1984.
- Frege, Gottlob *FTA* [1885]. Über formale Theorien der Arithmetik (On Formal Theories of Arithmetic). *Sitzungsberichte der Jenaischen Gesellschaft für Medizin und Naturwissenschaft* 19, suppl. 2, 94–104. Translation by E.-H. W. Kluge in Frege 1984, 112–121.
- Frege, Gottlob *GGZ* [1893–1903]. *Grundgesetze der Arithmetik, begriffsschriftlich abgeleitet*. Two volumes. Jena: Pohle. Partial translation in Frege 1964.
- Frege, Gottlob *LI* [1891]. Über das Trägheitsgesetz (On the Law of Inertia). *Zeitschrift für Philosophie und philosophische Kritik* 98, 145–61. Translation by Hans Kaal in Frege 1984.
- Frege, Gottlob *Neg* [1918]. Die Verneinung: Eine logische Untersuchung (Negation). *Beiträge zur Philosophie des deutschen Idealismus* 1, 143–57. Translated by Peter Geach and R. H. Stoothoff in Frege 1984.
- Frege, Gottlob *PW*. *Posthumous Writings*. Translation of Frege 1969 by Peter Long and Roger White. Chicago: University of Chicago Press, 1979.

- Frege, Gottlob *RH* [1894]. Review of E. Husserl, *Philosophie der Arithmetik*. *Zeitschrift für Philosophie und philosophische Kritik* 103, 313–32. Translated by Hans Kaal in Frege 1984.
- Frege, Gottlob *SJ* [1882]. Über die wissenschaftliche Berechtigung einer Begriffsschrift (On the Scientific Justification of a Conceptual Notation). *Zeitschrift für Philosophie und philosophische Kritik* 81, 29–32. Translation by T. W. Bynum in Frege 1972, 83–9. (Citation is to pagination in the translation.)
- Frege, Gottlob *SM* [1892]. Über Sinn und Bedeutung (On Sense and Meaning). *Zeitschrift für Philosophie und philosophische Kritik* 100, 25–50. Translation by Max Black in Frege 1984.
- Frege, Gottlob *Th* [1918]. Der Gedanke: Eine logische Untersuchung (Thoughts). *Beiträge zur Philosophie des deutschen Idealismus* 1, 58–77. Translated by Peter Geach and R. H. Stoothoff in Frege 1984.
- Friedman, Michael 1988. Logical Truth and Analyticity in Carnap's "Logical Syntax of Language." In Aspray and Kitcher 1988, 82–94.
- Friedman, Michael 1990. Form and Content. In Yourgrau 1990, 215–231.
- Friedman, Michael 1992. *Kant and the Exact Sciences*. Cambridge, MA: Harvard University Press.
- Fries, Jacob Friedrich 1837. *System der Logik*. In *Sämtliche Schriften*, bd. 7. Scientia Verlag Aalen, 1971.
- Gabbay, D. M., ed. 1994. *What Is a Logical System?* Oxford: Clarendon Press.
- Galen *IL*. *Institutio Logica*. (Introduction to Logic.) Ed. C. Kalbfleisch. Leipzig, 1896. Translation: Kieffer 1964.
- Garlandus Compotista [1040?] *D. Dialectica*. See De Rijk 1959.

- Gaukroger, Stephen 1989. *Cartesian Logic: An Essay on Descartes's Conception of Inference*. Oxford: Clarendon Press.
- Gentzen, Gerhard 1935. Untersuchungen über das logische Schliessen (Investigations into Logical Deduction). *Mathematisches Zeitschrift* 39, 176–210, 405–431. Trans. M. E. Szabo in *The Collected Papers of Gerhard Gentzen* (Amsterdam: North Holland, 1969), 68–131.
- Glouberman, M. 1979. Conceptuality: An Essay in Retrieval. *Kant-Studien* 70, 383–408.
- Gödel, Kurt 1953/9. Is Mathematics Syntax of Language? In Gödel 1995.
- Gödel, Kurt 1995. *Collected Works, Vol. III: Unpublished essays and lectures*. Ed. Solomon Feferman, John W. Dawson, Jr., Warren Goldfarb, Charles Parsons, and Robert M. Solovay. Oxford: Oxford University Press.
- Goldfarb, Warren 1979. Logic in the Twenties: The Nature of the Quantifier. *Journal of Symbolic Logic* 44, 351–368.
- Goldfarb, Warren 1988. Poincaré Against the Logicians. In Aspray and Kitcher 1988, 61–81.
- Goldfarb, Warren *FCL*. Frege's Conception of Logic. Forthcoming in *Future Pasts: Perspectives on the Place of the Analytic Tradition in Twentieth Century Philosophy*, ed. Juliet Floyd and Sanford Shieh.
- Gómez-Torrente, Mario 1996. Tarski on Logical Consequence. *Notre Dame Journal of Formal Logic* 37, 125–151.
- Goodman, Nelson 1961. About. *Mind* 70, 1–24. Reprinted in Goodman 1972, 246–272. (Citation is to reprint.)
- Goodman, Nelson 1972. *Problems and Projects*. Indianapolis: Bobbs-Merrill.

- Green-Pedersen, Niels Jørgen 1984. *The Tradition of the Topics in the Middle Ages: the Commentaries on Aristotle's and Boethius' 'Topics'*. Munich and Vienna: Philosophia Verlag.
- Gregory, Duncan 1838. On the Real Nature of Symbolical Algebra. *Edinburgh Philosophical Transactions* 14.
- Griffin, Nicholas 1993. Terms, Relations, Complexes. In Irvine and Wedeking 1993, 159–192.
- Guttenplan, Samuel, ed. 1975. *Mind and Language*. Oxford: Oxford University Press.
- Haack, Susan 1978. *Philosophy of Logics*. Cambridge: Cambridge University Press.
- Haaparanta, Leila, and Jaakko Hintikka, eds. 1986. *Frege Synthesized*. Dordrecht: Reidel.
- Hacking, Ian 1973. Leibniz and Descartes: Proof and Eternal Truths. *Proceedings of the British Academy* 59, 175–188.
- Hacking, Ian 1979. What Is Logic? *Journal of Philosophy* 76, 285–319.
- Hale, Bob 1999. On Some Arguments for the Necessity of Necessity. *Mind* 108, 23–52.
- Hamilton, William 1833. Recent Publications on Logical Science. *Edinburgh Review* 57, 194–238.
- Hamilton, William 1867. *Lectures on Metaphysics and Logic*. Vol. 2: Logic. Ed. Henry L. Mansel and John Veitch. Boston: Gould and Lincoln. (Lectures were originally composed in 1837–8.)
- Hanson, William H. 1997. The Concept of Logical Consequence. *Philosophical Review* 106, 365–409.
- Hanson, William H. and James Hawthorne 1985. Validity in Intensional Languages: A New Approach. *Notre Dame Journal of Formal Logic* 26, 9–35.

- Harman, Gilbert 1972. Is Modal Logic Logic? *Philosophia* 2, 75–84.
- Harman, Gilbert 1984. Logic and Reasoning. *Synthese* 60, 107–27.
- Hegel, Georg Wilhelm Friedrich 1827. *Encyclopaedia of the Philosophical Sciences in Outline*. Second ed. First part (Logic) translated in Wallace 1892. (Citation is to translation.)
- Hilbert, David 1899. *Grundlagen der Geometrie*. Leipzig: Teubner.
- Hintikka, Jaakko 1996. *The Principles of Mathematics Revisited*. Cambridge: Cambridge University Press.
- Hintikka, Jaakko 1998. *Language, Truth and Logic in Mathematics*. Dordrecht: Kluwer.
- Hintikka, Jaakko, et al., eds. 1978. *Essays in Mathematical and Philosophical Logic*. Boston: Reidel.
- Hughes, G. E. and Cresswell, M. J. 1968. *A New Introduction to Modal Logic*. London: Routledge.
- Hume, David *T* [1739]. *A Treatise of Human Nature*. Second ed. Ed. L. A. Selby-Bigge, rev. P. H. Nidditch. Oxford: Clarendon Press, 1978.
- Hylton, Peter 1984. The Nature of the Proposition and the Revolt Against Idealism. In Rorty, Schneewind, and Skinner 1984, 375–397.
- Hylton, Peter 1990a. Logic in Russell's Logicism. In Bell and Cooper 1990, 137–172.
- Hylton, Peter 1990b. *Russell, Idealism and the Emergence of Analytic Philosophy*. Oxford: Clarendon Press.
- Hyslop, James H. 1892. *The Elements of Logic: Theoretical and Practical*. New York: Charles Scribner's Sons.

- Irvine, A. D. and G. A. Wedeking, eds. 1993. *Russell and Analytic Philosophy*. Toronto: University of Toronto Press.
- Jané, Ignacio 1993. A Critical Appraisal of Second-order Logic. *History and Philosophy of Logic* 14, 67–86.
- Jevons, W. Stanley 1864. *Pure Logic or the Logic of Quality Apart from Quantity, with Remarks on Boole's System, and on the Relation of Logic to Mathematics*. In Jevons 1890. (Citation is to reprint.)
- Jevons, W. Stanley 1870. *Elementary Lessons in Logic*. London: Macmillan.
- Jevons, W. Stanley 1890. *Pure Logic and Other Minor Works*. Ed. Robert Adamson and Harriet A. Jevons. London: Macmillan.
- Kanger, S., and S. Öhman, eds. 1980. *Philosophy and Grammar*. Dordrecht: Reidel.
- Kant, Immanuel 1992. *Lectures on Logic*. Translated J. Michael Young. Cambridge: Cambridge University Press.
- Kant, Immanuel 1992b. *Theoretical Philosophy, 1755–1770*. Translated David Walford with Ralf Meerbote. Cambridge: Cambridge University Press.
- Kant, Immanuel *Ak*. *Kant's gesammelte Schriften*. 29 vols. Ed. Preussischen Akademie der Wissenschaften (I–XXII), Deutschen Akademie der Wissenschaften zu Berlin (XXIII), Akademie der Wissenschaften zu Göttingen (XXIV–XXIX). Berlin: Walter de Gruyter: 1902–.
- Kant, Immanuel *B* [1763]. *Der einzig mögliche Beweisgrund zu einer Demonstration des Daseins Gottes* (The only possible argument in support of a demonstration of the existence of God). In Kant *Ak:II*. Translated in Kant 1992b. (Citations are to the Academy pagination.)
- Kant, Immanuel *BL* [early 1770s]. *Blomberg Logic*. In Kant *Ak:XXIV*. Translated in Kant 1992.

- Kant, Immanuel *BuL* [early 1770s]. *Busolt Logic*. In Kant Ak:XXIV.
- Kant, Immanuel *D* [1764]. *Untersuchung über die Deutlichkeit der Grundsätze der natürlichen Theologie und der Moral* (Inquiry concerning the distinctness of the principles of natural theology and morality). In Kant Ak:II. Translated in Kant 1992b. (Citations are to the Academy pagination.)
- Kant, Immanuel *DWL* [1792]. *Dohna-Wundlacken Logic*. In Kant Ak:XXIV. Translated in Kant 1992.
- Kant, Immanuel *E* [1790]. *Über eine Entdeckung nach der alle neue Kritik der reinen Vernunft durch eine ältere entbehrlich gemacht werden soll*. In Kant 1923. Translation in Allison 1973.
- Kant, Immanuel *GMS* [1785]. *Grundlegung zur Metaphysik der Sitten*. In Kant Ak:IV. Translation: *Grounding for the Metaphysics of Morals*. Trans. James W. Ellington. Indianapolis: Hackett, 1981.
- Kant, Immanuel *ID* [1770]. *De Mundi Sensibilis atque Intelligibilis Forma et Principiis* (Inaugural Dissertation.) In Kant Ak:II. Translated in Kant 1992b. (Citations are to the Academy pagination.)
- Kant, Immanuel *JL*. *Jäsche Logic (Logic: A Manual for Lectures)*. Ed. G. B. Jäsche (1800). In Kant Ak:IX. Translated in Kant 1992. (Citations are to the Academy pagination.)
- Kant, Immanuel *KrV*. *Kritik der Reinen Vernunft*. Riga: Johann Friedrich Hartknoch. (First edition = A, 1781. Second edition = B, 1787.) In Kant Ak:IV. Translation: *Critique of Pure Reason*. Trans. Norman Kemp Smith. New York: St. Martin's, 1929.
- Kant, Immanuel *PC*. *Philosophical Correspondence, 1759–99*. Ed. and trans. Arnulf Zweig. Chicago: University of Chicago Press, 1967.

- Kant, Immanuel *PhL* [early 1770s]. *Philippi Logic*. In Kant Ak:XXIV.
- Kant, Immanuel *Pr* [1783]. *Prolegomena zu einer jeden künftigen Metaphysik, die als Wissenschaft wird auftreten können*. In Kant Ak:IV. Translation: *Prolegomena to Any Future Metaphysics*. Trans. Lewis White Beck, revising Carus. New York: Macmillan, 1950.
- Kant, Immanuel *PzL* [late 1780s]. *Pölitz Logic*. In Kant Ak:XXIV.
- Kant, Immanuel *R*. *Reflexionen*. (Handwritten notes on logic.) In Kant Ak:XVI, XVII.
- Kant, Immanuel *VL* [c. 1780]. *Vienna Logic*. In Kant XXIV. Translated in Kant 1992.
- Kaplan, David 1989 [1977]. Demonstratives: An Essay on the Semantics, Logic, Metaphysics, and Epistemology of Demonstratives and Other Indexicals. In Almog, Perry, and Wettstein 1989, 481–563.
- Keynes, John Neville 1906. *Studies and Exercises in Formal Logic, including a generalisation of logical processes in their application to complex inferences*. 4th ed. London: Macmillan.
- Kidd, I. G. 1978. Posidonius and Logic. In *Les Stoïciens et Leur Logique* (Actes du Colloque de Chantilly, 18–22 septembre 1976). Paris: J. Vrin.
- Kieffer, John Spangler 1964. *Galen's Institutio Logica*. (Translation, introduction, and commentary.) Baltimore: Johns Hopkins Press.
- Kitcher, Philip 1979. Frege's Epistemology. *Philosophical Review* 88, 235–262. Reprinted in Sluga ed. 1993.
- Klein, Felix 1893. Vergleichende Betrachtungen über neuere geometrische Forschungen (Das Erlanger Programm). [A Comparative Review of Recent Researches in Geometry.] *Mathematische Annalen* 43, 63–100.

- Kneale, William 1945–6. Truths of Logic. *Proceedings of the Aristotelian Society* n.s. 46, 206–234.
- Kneale, William 1956. The Province of Logic. In *Contemporary British Philosophy*. London: George Allen & Unwin Ltd.
- Kneale, William and Martha 1962. *The Development of Logic*. Oxford: Oxford University Press.
- Kremer, Michael 1988. Logic and Meaning: the Philosophical Significance of the Sequent Calculus. *Mind* 47, 50–72.
- Kretzmann, Norman, Anthony Kenny, and Jan Pinborg, eds. 1982. *The Cambridge History of Later Medieval Philosophy*. Cambridge: Cambridge University Press.
- Kuhn, Steven 1981. Logical Expressions, Constants, and Operator Logic. *Journal of Philosophy* 78 (1981), 487–499.
- Lambert, Johann Heinrich A. *Anlage zur Architektonic*. Riga, 1771. In *Philosophische Schriften*, vol. III, ed. Hans-Werner Arndt (Hildesheim: Georg Olms, 1965).
- Lambert, Karel 1970. *Philosophical Problems in Logic: Some Recent Developments*. Dordrecht: Reidel.
- Lee, Tae-Soo 1984. *Die griechische Tradition der aristotelischen Syllogistik in der Spätantike (Hypomnemata 79)*. Göttingen: Vandenhoeck and Ruprecht.
- Leibniz, G. W. *GM. Die Mathematische Schriften von Gottfried Wilhelm Leibniz*. 7 vols. Ed. C. J. Gerhardt: Berlin, 1849–55.
- Leibniz, G. W. *GP. Die Philosophischen Schriften von Gottfried Wilhelm Leibniz*. 7 vols. Ed. C. J. Gerhardt: Berlin, 1875–1890. Reprint Hildesheim: Georg Olms, 1960.
- Leibniz, G. W. *L. Philosophical Papers and Letters*. Ed. Leroy E. Loemker. Second edition. Dordrecht: D. Reidel, 1969.

- Leibniz, G. W. *NE. New Essays on Human Understanding*. Trans. and ed. Peter Remnant and Jonathan Bennett. Cambridge: Cambridge University Press, 1981.
- Lewis, Clarence Irving 1929. *Mind and the World Order: Outline of a Theory of Knowledge*. New York: Charles Scribner's Sons.
- Lewis, David 1970. General Semantics. *Synthese* 22, 18–67. Reprinted with postscripts in Lewis 1983. (Citations are to reprint.)
- Lewis, David 1983. *Philosophical Papers*. Vol. 1. Oxford: Oxford University Press.
- Lindenbaum, A. and A. Tarski 1936. On the Limitations of the Means of Expression of Deductive Theories. In Tarski 1983.
- Locke, John *E* [1700]. *An Essay Concerning Human Understanding*. Fourth ed. (London). Ed. Peter H. Nidditch. Oxford: Clarendon Press, 1975.
- Longuenesse, Béatrice 1998. *Kant and the Capacity to Judge: Sensibility and Discursivity in the Transcendental Analytic of the Critique of Pure Reason*. Trans. Charles T. Wolfe from French original (Presses Universitaires de France, 1993). Princeton: Princeton University Press.
- Lotze, Hermann 1843. *Logik*. Leipzig. Translation: Lotze 1888. (Citations are to translation.)
- Lotze, Hermann 1888. *Logic*. Trans. Bernard Bosanquet. Second ed. Vol. 1. Oxford: Clarendon Press.
- Lukasiewicz, Jan 1957. *Aristotle's Syllogistic*. Second ed. Oxford: Oxford University Press.
- Lycan, William G. 1989. Logical Constants and the Glory of Truth-Conditional Semantics. *Notre Dame Journal of Formal Logic* 30, 390–401.

- MacIntyre, Alasdair 1981. *After Virtue: A Study in Moral Theory*. Notre Dame, IN: University of Notre Dame Press.
- MacLane, Saunders 1996. Structure in Mathematics. *Philosophia Mathematica* 4, 174–183.
- MacLane, Saunders, and Garrett Birkhoff 1993. *Algebra*. Third edition. New York: Chelsea.
- Mansel, Henry Longueville 1851. *Prolegomena Logica: An Inquiry into the Psychological Character of Logical Processes*. Oxford: William Graham.
- Martin, Christopher 1986. William's Machine. *Journal of Philosophy* 83, 564–572.
- Massey, Gerald J. 1975. Are There Any Good Arguments That Bad Arguments Are Bad? *Philosophy in Context* 4, 61–77.
- Massey, Gerald J. 1976. Tom, Dick, and Harry, and All the King's Men. *American Philosophical Quarterly* 13, 89–107.
- Mates, Benson 1965. *Elementary Logic*. New York: Oxford University Press.
- Mautner, F. I. 1946. An Extension of Klein's Erlanger Program: Logic as Invariant-Theory. *American Journal of Mathematics* 68, 345–384.
- McCarthy, Timothy 1981. The Idea of a Logical Constant. *Journal of Philosophy* 78, 499–523.
- McDowell, John 1994. *Mind and World*. Cambridge, MA: Harvard University Press.
- McFetridge, I. G. 1990. Logical Necessity: Some Issues. In *Logical Necessity and Other Essays*, ed. John Haldane and Roger Scruton (Aristotelian Society Series, vol. 11, 1990), 135–54.
- McGee, Vann 1996. Logical Operations. *Journal of Philosophical Logic* 25, 567–580.

- Meier, Georg Friedrich A [1752]. *Auszug aus der Vernunftlehre*. Halle. Reprinted at the bottoms of pages (with Kant's marginalia) in Kant Ak:vol. 16.
- Mill, John Stuart 1889. *An Examination of Sir William Hamilton's Philosophy*. Sixth ed. London: Longmans, Green, and Co. First ed.: 1865.
- Mostowski, A. 1957. On a Generalization of Quantifiers. *Fundamenta Mathematicae* 44, 12–36.
- Mueller, Ian 1960. Stoic and Peripatetic Logic. In *Aristote et Les Problèmes de Méthode* (proceedings of Symposium Aristotelicum, Louvain, 1960), 2<sup>nd</sup> ed., Louvain: Editions de l'Institut Supérieur de Philosophie, 1980, 173–187.
- Mueller, Ian 1974. Greek Mathematics and Greek Logic. In Corcoran 1974, 35–70.
- Nagel, Ernest 1956. *Logic Without Metaphysics and Other Essays in the Philosophy of Science*. Glencoe, IL: Free Press.
- Nagel, Ernest 1979. *Teleology Revisited and Other Essays in the Philosophy and History of Science*. New York: Columbia University Press.
- Normore, Calvin G. 1993. The Necessity in Deduction: Cartesian Inference and Its Medieval Background. *Synthese* 96, 437–454.
- Owen, G. E. L. 1968, ed. *Aristotle on Dialectic: The Topics* (Proceedings of the Third Symposium Aristotelicum). Oxford: Clarendon Press.
- Parsons, Charles 1965. Frege's Theory of Number. In *Philosophy in America*, ed. Max Black. Ithaca: Cornell University Press, 180–203. Reprinted in Parsons 1983, 150–175.
- Parsons, Charles 1969. Kant's Philosophy of Arithmetic. In *Philosophy, Science, and Method: Essays in Honor of Ernest Nagel*, ed. Sidney Morgenbesser, Patrick Suppes, and Morton White. New York: St. Martin's Press. Reprinted with postscript in Parsons 1983, 110–149. Citations to reprint.

- Parsons, Charles 1983. *Mathematics in Philosophy*. Ithaca: Cornell University Press.
- Pasch, M. 1926. *Vorlesungen über Neuere Geometrie*. Second edition. (First edition, 1882). Berlin: Verlag von Julius Springer.
- Paton, H. J. 1936. *Kant's Metaphysic of Experience*. 2 vols. New York: Macmillan, 1936.
- Peacocke, Christopher 1976. What Is a Logical Constant? *Journal of Philosophy* 73, 221–240.
- Peacocke, Christopher 1987. Understanding Logical Constants: A Realist's Account. *Proceedings of the British Academy* 73, 153–200.
- Philoponus, John *AnPr*. In *Aristotelis Analytica Priora Commentaria*. Ed. M. Wallies. (*Commentaria in Aristotelem Graeca* XIII 2). G. Reimer, 1905.
- Plotinus *Enn. Enneads*. Six volumes. Trans. A. H. Armstrong. Cambridge, MA: Harvard University Press, 1966.
- Poincaré, Henri 1902. *La Science et l'Hypothèse*. Paris: Flammarion. Translated in Poincaré 1946.
- Poincaré, Henri 1908. *Science et Méthode*. Paris: Flammarion. Translated in Poincaré 1946.
- Poincaré, Henri 1946. *The Foundations of Science*. Trans. George Bruce Halsted. Lancaster, PA: The Science Press.
- Pollock, John 1967. Logical Validity in Modal Logic. *The Monist* 51, 128–135.
- Popper, Karl 1947. New Foundations for Logic. *Mind* 56, 193–235.
- Prawitz, Dag 1978. Proofs and the Meaning and Completeness of the Logical Constants. In Hintikka et al. 1978, 25–39.
- Prior, A. 1960. The Runabout Inference-Ticket. *Analysis* 21, 38–9.

- Prior, A. 1976. *Papers in Logic and Ethics*. Ed. P. T. Geach and A. J. P. Kenny. Amherst, MA: University of Massachusetts Press.
- Pseudo-Alexander [Michael of Ephesus] *SE. Alexandri quod fertur in Aristotelis Sophisticos Elenchos Commentarium (Commentaria in Aristotelem Graeca II 3)*. Ed. M. Wallies. Berlin: G. Reimeri, 1898.
- Pseudo-Ammonius *AnPr. In Aristotelis Analyticorum Priorum Librum I Commentarium (Commentaria in Aristotelem Graeca IV 6)*. Ed. M. Wallies. Berlin: G. Reimeri, 1899.
- Putnam, Hilary 1975. The Meaning of "Meaning." In *Language, Mind, and Knowledge* (Minnesota Studies in the Philosophy of Science VII), ed. K. Gunderson (Minneapolis: University of Minnesota Press).
- Putnam, Hilary 1969. Is Logic Empirical? In *Boston Studies in the Philosophy of Science V*, ed. Robert S. Cohen and Marx W. Wartofsky (Dordrecht: D. Reidel, 1969), 216–241.
- Putnam, Hilary 1994a. *Words and Life*. Ed. James Conant. Cambridge, MA: Harvard University Press.
- Putnam, Hilary 1994b. Rethinking Mathematical Necessity. In Putnam 1994a, 245–63.
- Quine, W. V. 1963. Carnap and Logical Truth. In Schilpp 1963. Reprinted in Quine 1976, 106–132.
- Quine, W. V. 1976. *The Ways of Paradox and Other Essays*. Revised and enlarged edition. Cambridge, MA: Harvard University Press.
- Quine, W. V. 1980. Grammar, Truth, and Logic. In Kanger and Öhman 1980, 17–28.
- Quine, W. V. 1986. *Philosophy of Logic*. Second edition. Cambridge, MA: Harvard University Press.

- Resnik, Michael 1988. Second-order Logic Still Wild. *Journal of Philosophy* 85, 75–87.
- Ricketts, Thomas 1985. Frege, the *Tractatus*, and the Logocentric Predicament. *Nous* 19, 3–15.
- Ricketts, Thomas 1986. Objectivity and Objecthood: Frege's Metaphysics of Judgment. In Haaparanta and Hintikka 1986, 65–95.
- Ricketts, Thomas 1996. Logic and Truth in Frege. *Proceedings of the Aristotelian Society* s. v. 70, 121–40.
- Ricketts, Thomas 1997. Frege's 1906 Foray into Metalogic. *Philosophical Topics* 25, 169–188.
- Rorty, R., J. Schneewind, and Q. Skinner 1984. *Philosophy in History*. Cambridge: Cambridge University Press.
- Ross, W. D., ed. 1949. *Aristotle's Prior and Posterior Analytics*. Oxford: Clarendon Press.
- Ross, W. D., ed. 1958. *Aristotelis Topica et Sophistici Elenchi*. Oxford: Clarendon Press.
- Russell, Bertrand 1900. *A Critical Exposition of the Philosophy of Leibniz*. London: George Allen and Unwin, new ed. 1937.
- Russell, Bertrand 1903. *The Principles of Mathematics*. First ed. Cambridge: Cambridge University Press.
- Russell, Bertrand 1911. The Philosophical Implications of Mathematical Logic. *Revue de Métaphysique et de Morale* 19, 281–91. Translation by P. E. B. Jourdain with revisions by Russell in *The Monist* 22 (1913), 481–93. Reprinted in Russell 1973, 284–294. (Citation is to reprint.)
- Russell, Bertrand 1913. *Theory of Knowledge: The 1913 Manuscript*. London: George Allen and Unwin, 1984. Citations are to paperback ed., London: Routledge, 1992.

- Russell, Bertrand 1914. *Our Knowledge of the External World as a Field for Scientific Method in Philosophy*. London: George Allen and Unwin.
- Russell, Bertrand 1920. *Introduction to Mathematical Philosophy*. Second edition. London: George Allen and Unwin.
- Russell, Bertrand 1938. *The Principles of Mathematics*. Second ed. New York: Norton.
- Russell, Bertrand 1973. *Essays in Analysis*. London: Allen and Unwin.
- Ryle, Gilbert 1954. *Dilemmas*. Cambridge: Cambridge University Press.
- Sainsbury, Mark 1991. *Logical Forms: An Introduction to Philosophical Logic*. Oxford: Basil Blackwell.
- Sánchez-Miguel, Manuel García-Carpintero 1993. The Grounds for the Model-theoretic Account of the Logical Properties. *Notre Dame Journal of Formal Logic* 34, 107–131.
- Schilpp, P. A., ed. 1963. *The Philosophy of Rudolf Carnap*. La Salle, IL: Open Court.
- Schirn, Matthias 1998. *The Philosophy of Mathematics Today*. Oxford: Clarendon Press.
- Schlick, M. 1925. *Allgemeine Erkenntnislehre*. Second ed. Berlin: Springer. Citations are to the translation, Schlick 1974.
- Schlick, M. 1974. *General Theory of Knowledge*. Trans. A. E. Blumberg. New York: Springer.
- Schmidt, Robert W., S. J. 1966. *The Domain of Logic According to Saint Thomas Aquinas*. The Hague: Martinus Nijhoff.
- Schroeder-Heister, Peter 1984. Popper's Theory of Deductive Inference and the Concept of a Logical Constant. *History and Philosophy of Logic* 5, 79–110
- Scott, Dana 1970. Advice on Modal Logic. In Lambert 1970, 143–173.
- Sellars, Wilfrid 1953. Inference and Meaning. *Mind* 62, 313–338.

- Sellars, Wilfrid 1967. *Science and Metaphysics: Variations on Kantian Themes*. (John Locke Lectures, 1965–6.) London: Routledge and Kegan Paul. Reprint: Atascadero, CA: Ridgeview Publishing Company, 1992.
- Sextus Empiricus *PH. Outlines of Pyrrhonism*. Greek text with trans. by R. G. Bury. Cambridge, MA: Harvard University Press, 1933.
- Shapiro, Stewart 1991. *Foundations Without Foundationalism: A Case for Second-Order Logic*. Oxford: Oxford University Press.
- Shapiro, Stewart 1993. Modality and Ontology. *Mind* 102, 455–481.
- Shapiro, Stewart 1997. *Philosophy of Mathematics: Structure and Ontology*. Oxford: Oxford University Press.
- Shapiro, Stewart 1998a. Proof and Truth: Through Thick and Thin. *Journal of Philosophy* 95, 493–521.
- Shapiro, Stewart 1998b. Logical Consequence: Models and Modality. In Schirn 1998, 131–156.
- Sher, Gila 1991. *The Bounds of Logic: a Generalized Viewpoint*. Cambridge, MA: MIT Press.
- Sher, Gila 1996. Did Tarski Commit “Tarski’s Fallacy”? *Journal of Symbolic Logic* 61, 653–686.
- Simons, Peter 1998. Structure and Abstraction. In Schirn 1998.
- Simmons, Edward D. 1961. The Nature and Limits of Logic. *The Thomist* 24, 47–71.
- Sluga, Hans 1976. Frege as a Rationalist. In *Studien zu Frege I: Logik und Philosophie der Mathematik*, ed. Matthias Schirn (Stuttgart: Frommann-Holzboog), 27–47. Reprinted in Sluga 1993.
- Sluga, Hans 1980. *Gottlob Frege*. London: Routledge.

- Sluga, Hans 1993, ed. *The Philosophy of Frege, Volume 1: General Assessments and Historical Accounts of Frege's Philosophy* (New York: Garland).
- Stalnaker, Robert C. 1970. Pragmatics. *Synthese* 22. Citation is to reprint in Stalnaker 1999.
- Stalnaker Robert C. 1999. *Context and Content*. Oxford: Oxford University Press.
- Strawson, P. F. 1952. *Introduction to Logical Theory*. London: Methuen.
- Stump, Eleonore 1978. *Boethius's De Topicis Differentiis*. (Translation with notes and essays.) Ithaca, NY: Cornell University Press.
- Stump, Eleonore 1982. Topics: Their Development and Absorption into Consequences. In Kretzmann et al. 1982.
- Stump, Eleonore 1988. *Boethius's In Ciceronis Topica*. Trans. with notes. Ithaca, NY: Cornell University Press.
- Stump, Eleonore 1989. *Dialectic and Its Place in the Development of Medieval Logic*. Ithaca, NY: Cornell University Press.
- Tappenden, Jamie 1995. Geometry and Generality in Frege's Philosophy of Arithmetic. *Synthese* 102, 319–361.
- Tappenden, Jamie 1997. Metatheory and Mathematical Practice in Frege. *Philosophical Topics* 25, 213–264.
- Tarski, Alfred 1936. On The Concept of Logical Consequence. In Tarski 1983, 409–420.
- Tarski, Alfred 1944. A Philosophical Letter of Alfred Tarski. *Journal of Philosophy* 84 (1987), 28–32.
- Tarski, Alfred 1965. *Introduction to Logic and to the Methodology of Deductive Sciences*. New York: Oxford University Press.

- Tarski, Alfred 1966. What are Logical Notions? Ed. John Corcoran. *History and Philosophy of Logic* 7 (1986), 143–154
- Tarski, Alfred 1983. *Logic, Semantics, Metamathematics*. Second edition. Trans. J. H. Woodger, ed. John Corcoran. Indianapolis: Hackett, 1983.
- Tennant, Neil 1997. On the Necessary Existence of Numbers. *Nous* 31, 307–336.
- Tharp, Leslie H. 1975. Which Logic Is the Right Logic? *Synthese* 31, 1–21.
- Thomason, Richmond H. 1984. Combinations of Tense and Modality. In *Handbook of Philosophical Logic*, Vol. 2, ed. D. Gabbay and F. Guenther. Dordrecht: Reidel.
- Thomson, William 1860. *An Outline of the Necessary Laws of Thought: A Treatise on Pure and Applied Logic*. Fifth ed. London: Longman, Green, Longman, and Roberts.
- Thompson, Manley 1972–3. Singular Terms and Intuitions in Kant's Epistemology. *Review of Metaphysics* 26, 314–343.
- Trendelenburg, Adolf 1867. Über Leibnizes Entwurf einer allgemeinen Charakteristik (On Leibniz's Project of a Universal Characteristic). In *Historische Beiträge zur Philosophie* 3. Berlin.
- Trendelenburg, Adolf 1870. *Logische Untersuchungen*. Hildesheim: Georg Olms, 1964. Reprint of 3rd edition, Leipzig, 1870.
- Ueberweg, Friedrich 1882. *System der Logik und Geschichte der Logischen Lehren*. Bonn: Adolph Marcus.
- van Benthem, Johan 1989. Logical Constants Across Varying Types. *Notre Dame Journal of Formal Logic* 30, 315–342.
- van Heijenoort, Jean 1967. Logic as Calculus and Logic as Language. *Synthese* 17.
- van Heijenoort, Jean, ed. 1967. *From Frege to Gödel*. Cambridge, MA: Harvard University Press.

- Wagner, Steven J. 1987. The Rationalist Conception of Logic. *Notre Dame Journal of Formal Logic* 28, 3–35.
- Wallace, William 1892. *The Logic of Hegel*. (Translation of first part of Hegel 1827.) Oxford: Clarendon Press.
- Wang, Hao 1996. *A Logical Journey: From Gödel to Philosophy*. Cambridge, MA: MIT Press.
- Warmbröd, Ken 1999. Logical Constants. *Mind* 108, 503–38.
- Weiner, Joan 1986. Putting Frege in Perspective. In Haaparanta and Hintikka 1986, 9–30.
- Weiner, Joan 1990. *Frege in Perspective*. Ithaca: Cornell University Press.
- Whately, Richard 1826. *Elements of Logic*. London: J. Mawman. Reprint, ed. Paola Dessì, Bologna: Cooperativa Libreria Universitaria Editrice, 1988.
- Wittgenstein, Ludwig 1979. *Notebooks 1914–1916*. Second edition. Ed. G. H. von Wright and G. E. M. Anscombe, trans. G. E. M. Anscombe. Chicago: University of Chicago Press.
- Wittgenstein, Ludwig 1922. *Tractatus Logico-Philosophicus (Logisch-Philosophische Abhandlung)*. Trans. C. K. Ogden. London: Routledge.
- Wittgenstein, Ludwig 1953. *Philosophical Investigations*. Trans. G. E. M. Anscombe. Oxford: Blackwell.
- Wolff, Christian L [1728]. *Philosophia Rationalis Sive Logica*. Third ed., Frankfurt. In *Christian Wolff: Gesammelte Werke*, II.1, ed. J. École et al. (Hildesheim and New York: Georg Olms, 1983).
- Wright, Crispin 1983. *Frege's Conception of Numbers as Objects*. Aberdeen: Aberdeen University Press (Scots Philosophical Monographs).
- Wright, Crispin 1998. On the Harmless Impredicativity of  $N^=$ . In Schirn ed. 1998.

Yourgrau, Paul, ed. 1990. *Demonstratives*. Oxford: Oxford University Press.