The Birth of Semantics

Richard Kimberly Heck and Robert May

Brown University and University of California, Davis

Abstract

We attempt here to trace the evolution of Frege's thought about truth. What most frames the way we approach the problem is a recognition that hardly any of Frege's most familiar claims about truth appear in his earliest work. We argue that Frege's mature views about truth emerge from a fundamental re-thinking of the nature of logic instigated, in large part, by a sustained engagement with the work of George Boole and his followers, after the publication of *Begriffsschrift* and the appearance of critical reviews by members of the Boolean school.

There could hardly be more disagreement about the role that the notion of truth does or does not play in Gottlob Frege's philosophy. According to the influential interpretation developed by Sir Michael Dummett (1981), Frege's contribution to modern logic was not simply the introduction of the notion of quantification and, with it, a formal language adequate for the purposes of mathematics. Frege also provided a semantic theory for his formal language, one that was similar in spirit (and in places in detail) to the theories of truth developed by Alfred Tarski (1956) in the 1930s. Such now familiar doctrines as that the meaning of a sentence is its truth-condition are central to Frege's work and justify our regarding him as the founding father (or, at least, grandfather) of formal semantics and philosophy of language, as well as of mathematical logic. On the other side, defenders of the so-called universalist interpretation of Frege, such as Thomas Ricketts (1997), hold that deep features of Frege's philosophy precluded him from even attempting to give a semantic theory for his formal language. In a similar vein, many philosophers have supposed that Frege is some sort of deflationist, in light of his remark in "On Sense and Reference" that "the sentence 'The thought that 5 is a prime number is true' contains. . . the same thought as the simple '5 is a prime number' " (SM, op. 34; cf. MBLI, p. 251).

Our sympathies are with Dummett, and our goal here is to defend, extend, and amend his interpretation. We will begin by arguing that the texts usually taken to support the deflationary interpretation do not really do so.[1] But we think there are other texts that do support that sort of reading. That is because the semantic perspective that Dummett finds in Frege's mature philosophy is almost wholly absent from his earliest work, in particular, from *Begriffsschrift*. Indeed, as we shall see, semantic notions do not appear explicitly in Frege's corpus until the early 1890s, though they seem to have been implicit in his work as early as 1884. The interesting question, historically, is why Frege's thinking took this 'semantic turn'. But the question is not just of historical interest. Understanding why Frege took the semantic turn can only contribute to our understanding of what is involved in our continuing to do so.

1 The Regress Argument

Much of the existing discussion of Frege's views on truth focuses on an argument he gives for the conclusion that truth is indefinable. This argument, which has come to be known as the 'regress argument', appears in at least two places: the late essay, "Thoughts", written around 1917 (Tht, op. 60), and an unfinished essay, "Logic", which the editors of the *Nachlass* date to 1897. Here is the argument as it occurs in "Logic":[2]

> Now it would be futile to employ a definition in order to make it clearer what is to be understood by 'true'. If, for example, we wished to say that 'an idea is true if it agrees with reality', nothing would have been achieved, since in order to apply this definition we should have to decide whether some idea or other did agree with reality[, in other words: whether it is true that the idea agreed with reality].[3] Thus we should have to presuppose the very thing that is being defined. The same would hold good of any definition of the form 'A is true if and only if it has such-and-such properties or stands in such-and-such a relation to such-and-such a thing'. In each case in

[1] For a bit more on the question whether Frege was a deflationist, see (Heck and May, 2018, §6).

[2] Dirk Greimann (2015) offers a very different interpretation of this passage.

[3] The bracketed words are missing from the English translation but are present in the German version, at least as printed in *Nachgalessene Schriften*. Thanks to Marcus Rossberg for drawing this to our attention.

> hand it would always come back to the question whether it is true that A has such-and-such properties, stands in such-and-such a relation to such-and-such a thing. Truth is obviously something so primitive and simple that it is not possible to reduce it to anything still simpler. (Log97, pp. 128–9)

Some commentators have found in this passage an argument that there is no real property of truth at all (Kemp, 1995; Ricketts, 1996, §II), which of course fuels the deflationary reading of Frege. But we do not think the argument supports that conclusion.

The regress argument is extremely puzzling. Frege twice insists that it is not enough to ask whether an idea agrees with reality. We must also ask whether it is *true* that the idea agrees with reality. But that seems gratuitous. Surely it is possible to ask whether 5 is prime without asking whether it is true that 5 is prime, let alone whether it is true that it is true that 5 is prime, and so on *ad infinitum*. Indeed, if it is not, then there looks to be a threat of regress whether truth is definable or not.

What is supposed to grease these apparent slides is revealed by remarks that follow the regress argument proper:

> What, in the first place, distinguishes ['true'] from all other predicates is that predicating it is always included in predicating anything whatever. If I assert that the sum of 2 and 3 is 5, then I thereby assert that it is true that 2 and 3 make 5. So I assert that it is true that my idea of Cologne Cathedral agrees with reality, if I assert that it agrees with reality. Therefore, it is really by using the form of an assertoric sentence that we assert truth, and to do this we do not need the word 'true'. (Log97, p. 129)

Frege is thus claiming that every assertion is, of its very nature, an assertion of truth. And every judgement is a judgement of truth.

Frege makes this sort of claim in other places, too, for example, in "On Sense and Reference", where he writes: "A judgement, for me, is not the mere grasping of a thought, but the admission of its truth" (SM, op. 34, note). This view is central to Frege's mature conception of logic. In many of his writings from just after the publication of *Begriffsschrift*, Frege emphasizes that he "did not wish to present an abstract logic in formulas, but to express a content through written symbols in a more precise and perspicuous way than is possible with words" (AimCN, pp. 90–1). Frege's logic was to be one we can actually use in reasoning, in proving theorems, where theorems are true contents. So logic issues in judgements.

It is important to appreciate, however, that the judgements in which logic issues need not themselves be logical truths. In Frege's own work, of course, he proves only theorems of logic, since he is particularly concerned with the question how much mathematics is implicit in logic itself. But Frege clearly anticipated and expected that his logic would eventually be expanded to include parts of 'science' that were not logical in character, for example, geometry. In principle, that is, Frege would have been perfectly happy to add the axioms of Euclidean geometry to his system as basic, underived principles from which proofs might begin. These would not be logical principles, so the theorems proven from them would not (typically) be logical theorems. But the very same logical inferences and Basic Laws that permit us to prove, say, the transitivity of the ancestral are also what permit us to derive Euclid's results from the axioms of geometry. That is the sense in which logic is supposed to be universal, or topic-neutral (May, 2018).

Frege's insistence that logic issues in judgements should not, then, be taken to limit the scope of logic to the derivation of logical truths. Rather, Frege's point is that logic *operates on* judgements (Geo2, op. 387). Inferences, as Frege understands them, are not just transitions between thoughts but transitions between judgements. But not just any transition between judgements counts as an inference, let alone as a valid one—the sort of inference of which logic must take notice. Logic is interested only in inferences that are "so constituted that, if a new proposition is derived from true propositions in accordance with them, it too is true" (Gg, v. II, §104). Logic, that is to say, is interested only in rules of inference that are truth-preserving.

But why is logic interested only in transitions between judgements that are truth-preserving? The answer, for Frege, is connected with logic's ambition to contribute to the growth of knowledge: "Knowledge must stand as the goal", he tells us, "and everything that occurs must be determined thereby" (Gg, v. II, §92). Indeed, Frege seems to think that the transition from, e.g., "$8 > 6$" and "$3 + 5 = 8$" to "$3 + 5 > 6$" is licensed *only in so far* as the goal is knowledge (Gg, v. II, §104). If our goal were to create pretty wallpaper, for example, then perhaps that 'transition' would not be allowed. But if our goal is knowledge, then it is, because the inference is guaranteed to be truth-preserving.

To put it differently, Frege thinks of judgement as having an aim that logic helps us achieve. Judgement aims at knowledge and therefore at truth, and reasoning in accord with the principles of logic guards one against straying from the path to truth. Frege's logic improves on what preceded it not so much by articulating modes of reasoning that are guaranteed to be truth-preserving—both Aristotle and George Boole managed that much—but by providing us with a rigorous way of determining whether some chain of reasoning has in fact restricted itself to those reliable modes. But here again, reasoning means: transitions between judgements. And the truth-preserving transitions are of special interest only because judgement has a peculiarly intimate relation to truth. It is this intimate relationship between judgement and truth that Frege is trying to articulate when he says that every judgement is a judgement of truth.

This view, that "predicating [truth] is always included in predicating anything whatever" (Log97, p. 129), seems to be one that Frege held throughout his career. But he understood this thesis in different ways at different times. Consider the following passage from *Begriffsschrift*:

> We can imagine a language in which the proposition "Archimedes perished at the capture of Syracuse" would be expressed thus: "The violent death of Archimedes at the capture of Syracuse is a fact". To be sure, one can distinguish between subject and predicate here, too, if one wishes to do so, but the subject contains the whole content, and the predicate serves only to turn the content into a judgement. *Such a language would have only a single predicate for all judgements, namely, "is a fact". . . . Our begriffsschrift is a language of this sort, and in it the sign $\vdash$ is the common predicate for all judgements*. (Bg, §3, emphasis original)

These remarks come at the conclusion of Frege's explanation of why the "distinction between subject and predicate does not occur in [his] way of representing a judgement" (Bg, §3). It is tempting, therefore, to regard them as but a grudging concession to tradition. But note how Frege emphasizes the final sentence of the quoted passage. This is something he does throughout Part I of *Begriffsschrift* when he is articulating the central features of his new conception of logic.[4] Frege is thus saying here, as explicitly and emphatically as he can, that it is one of the characteristic features of his formal language that, in it, there is only one (real) predicate: the assertion-sign.

[4] One can verify this fact by looking through Part I. Such emphasized remarks appear, for example, at the end of §1, §2, §3 (the one we just quoted), and §4.

What is perhaps most striking is Frege's claim that "the predicate ['is a fact'] serves. . . to turn [a] content into a judgement". This language is reminiscent of remarks Frege makes about the assertion-sign earlier in *Begriffsschrift*: When that sign is absent, we have "a mere combination of ideas", but a "content becomes a judgement when $\vdash$ is written before its sign. . . " (Bg, §2). So Frege is telling us that assertion is achieved through the predication of facthood, and that this is expressed in his new logic by the assertion-sign.[5]

This very idea is explicitly criticized in Frege's later work—though, as often, Frege does not own up to the fact that he is criticising his own earlier view. Frege writes in "On Sense and Reference":

> One might be tempted to regard the relation of the thought to the True not as that of sense to reference, but rather as that of subject to predicate. One can indeed say: "The thought that 5 is a prime number is true". But closer examination shows that nothing more has been said than in the simple sentence "5 is a prime number". The truth claim arises in each case from the form of the assertoric sentence, and when the latter lacks its usual force, e.g., in the mouth of an actor upon the stage, even the sentence "The thought that 5 is a prime number is true" contains only a thought, and indeed the same thought as the simple "5 is a prime number". It follows that the relation of the thought to the True may not be compared with that of subject to predicate. (SM, op. 34)

As we have already mentioned, Frege's claim that the attribution of truth adds nothing to the sense is often interpreted as an expression of a deflationary attitude towards truth. But Frege's central point here does not depend upon that claim, which arguably conflicts with his other views about sense.[6] Frege's main point in this passage is that one cannot

[5] As Ian Proops (1997) points out, Ludwig Wittgenstein ascribes this view to Frege: "The verb of a proposition is not 'is true' or 'is false', as Frege thought; rather, that which 'is true' must already contain the verb" (Wittgenstein, 1961, 4.063; see also Wittgenstein, 1979a, pp. 93, 100). Proops was the first to notice Frege's commitment to this view in *Begriffsschrift*.

[6] There are several places in Frege's writings where he makes claims of sameness of sense that one might think he really ought not make, e.g., that "$A \wedge B$" has the same sense as "$B \wedge A$" (CT, op. 39). Frege is led to make such claims, we would argue, by an incorrect application of his famous criterion for difference of sense: that one should be able to believe p but not q. As Frege usually applies that criterion, it acts as a *sufficient* condition for difference of sense. Sometimes, however, he seems to apply it as if it were a

assert a thought just by predicating truth of it. Whether doing so changes what thought is expressed is irrelevant. Even if "The thought that 5 is a prime number is true" and "5 is a prime number" express different thoughts, the former no more contains a judgement than does the latter. Frege mentions the theater to make this point, but he might equally have mentioned conditionals, such as: If the thought that 5 is a prime number is true, then there are odd primes. Frege famously emphasizes elsewhere (e.g. BLC, p. 11; IntLog, pp. 185–6) that the antecedents of conditionals are not, in any sense, asserted, and that remains the case even when the word 'true' appears in the antecedent.

Just because one has predicated truth of a thought, then, one has not necessarily made the judgement with that thought as its content. But there is, of course, another sense of 'predicate', in which one does not predicate whiteness of snow if one merely supposes that snow is white, but does predicate whiteness of snow if one judges that snow is white. So one might suggest that it is only in this other sense that judging involves predicating truth of a thought. What the regress argument shows is that this cannot be right. Predication in this other sense *is itself a sort of judgement*, at least according to Frege (IntLog, p. 185): If so, however, then to predicate truth of the thought that p is just to judge that the thought that p is true, that is, to judge that it is true that p. If so, however, then judging that it is true that p is predicating truth of the thought that it is true that p, that is, judging that it is true that it is true that p. And now we really *have* started a regress. Moreover, the regress is vicious, since the sense in which judgement *is* predication of truth was meant to be constitutive.[7] So it is not just that to "assert truth... we do not need the word 'true'", as Frege puts it in "Logic".[8] The

necessary condition, but it has no plausibility as a necessary condition, and taking it to be one conflicts with other of Frege's views about sense (Heck and May, 2010, §5, esp. n. 72). That is true here, too. The sentence "The thought that 5 is a prime number is true" refers to a thought, just as "John believes that 5 is a prime number" does, whereas "5 is a prime number" does not refer to a thought but only to a number (cf. Moore, 1953, p. 276). So they cannot have the same sense.

[7] It would not be vicious if the claim were, say, merely one about the commitments one incurs by making a judgement, as Dummett (1981, ch. 13) makes clear.

[8] In a brief note composed in 1915, we find Frege writing, in a similar vein: "... [A]ssertion is not to be found in the word 'true', but in the assertoric force with which the sentence is uttered" (MBLI, p. 251). (The surrounding material is again concerned with the consequences of the alleged fact that "Sea-water is salt" and "It is true that sea-water is salt" have the same sense.)

right conclusion to draw is that judgement is not predication of truth, in any sense.

In place of that view, Frege offers his mature view that "the relation of the thought to the True [is] that of sense to reference" (SM, op. 34). Frege thus does not abandon the idea that there is an intimate relationship between judgement and truth. Rather, he reconceives the nature of that relationship. He originally thought that judgement was predication of truth (or facthood). But the regress argument refutes that view: Judgement might in some sense involve the acceptance of thoughts 'as true', but we cannot understand what that means in terms of a thinker's predicating truth of a thought. Fundamentally, then, truth is not a property of thoughts: The most fundamental relation in which a thought stands to its truth-value is not the relation of subject to predicate but that of sense to reference.

One might well say, therefore, that Frege's objection to the predicational view is that it does not give truth a role that is central enough. Judging that 5 is prime has to involve something more than merely entertaining the thought that 5 is prime, but what more judgement adds cannot be understood in terms of our predicating truth of this thought. The problem is that predication happens at the level of the thought: The thought that p is true is just another thought (and again, it does not really matter whether it is the same thought or a different one). What we do in judgement, by contrast, is to take "the step from the level of thoughts to the level of reference" (SM, op. 34), and that is where the True and the False are properly to be located.[9]

This is, we think, a promising and underappreciated idea, quite at odds with the way philosophers nowadays tend to think about truth.[10] At least since Tarski (1956), philosophical discussions of truth have tended to focus on the so-called 'truth-predicate'. What Frege is telling us is that

[9] It does not follow that there is no property had by all and only thoughts that refer to the True. Indeed, we have just said which property that is. Still, it is easy to understand why some commentators have been tempted to read Frege as arguing for the stronger conclusion that there is no property of truth at all. To read him so, however, is to miss what is really at issue, which is a question about how we should understand the relation between a thought and its truth-value. There is such a relation, and so there is such a property, but that property is not what is fundamental. Rather, it is derivative from the relation between a thought and its truth-value.

[10] Something like this view is also to be found in Donald Davidson (1984; 1990; 2005), Dummett (1978; 1991, chs. 1–6; 1993b), and David Wiggins (1980), among others. Dorit Bar-On and Keith Simmons (2007), Mark Textor (2010), and Walter B. Pedriali (2017) discuss this aspect of Frege's view, as does Dummett (1981, esp. chs. 10 and 12).

this focus is misplaced. The study of truth must not be confused with the study of how the word “true” is or should be used in natural language. The notion of truth must already be in place before we can even begin to discuss the use of “true”, because a distinction between truth and falsity is implicit in the act of judging itself. Such a distinction must, as Frege puts it, be “recognized. . . by everybody who judges something to be true—and so even by a skeptic” (SM, op. 34).

A similar point is expressed by Sir Peter Strawson when he writes:[11]

> The occurrence in ordinary discourse of the words “fact”, “statement”, “true” signalizes the occurrence of [fact-stating] discourse. . . . If our task were to elucidate the nature of [this] type of discourse, it would be futile to attempt to do so in terms of the words “fact”, “statement”, “true”, for those words contain the problem not its solution. It would, for the same reason, be equally futile to attempt to elucidate any one of these words (in so far as the elucidation of *that* word would be the elucidation of *this* problem) in terms of the others. And it is, indeed, very strange that people have so often proceeded by saying, “Well, we’re pretty clear about what a statement is, aren’t we? Now let us settle *the further question*, *viz*., what it is for a statement to be true.” This is like “Well, we’re clear about what a command is: now what is it for a command to be obeyed?” (Strawson, 1950, p. 141, emphasis original, punctuation corrected)

And it is very strange. Or, to put it differently, it should not be at all surprising that, *if* you are prepared to assume that you know what statements are, *then* you should find that there is nothing left to say about what it is for a statement to be true. But that is not because there is nothing substantial to be said about truth. It is because what there is to be said about it must be said in the course of ‘elucidating’ the nature

[11] This remark, we suggest, makes it far less clear than it is usually supposed to be that Strawson is any kind of deflationist. He thinks there is little interesting to be said about the use of the word “true”, but he thinks there is a different problem, that of “elucidating the fact-stating type of discourse” (Strawson, 1950, p. 142), that is well worth pursuing. That, according to Strawson, is where the philosophical puzzles about truth really lie, even though Strawson himself reserves the label “the problem of truth” for questions about “true”. But that is just a consequence of the intellectual culture in which Strawson was working.

of judgement or, perhaps, the contents of judgements, which is what Strawson means by a ‘statement’.[12]

Our goal here, however, is not to try to develop Frege’s idea that truth and falsity are the references of sentences. Rather, in the remainder of this paper, we want to try to understand the origins of this now negelected view.

2 The Truth-Value Thesis

According to Frege, truth is the real subject-matter of logic. As he puts it in a now famous passage:[13]

> Just as ‘beautiful’ points the way for aesthetics and ‘good’ for ethics, so does the word ‘true’ for logic. 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. (Tht, op. 58)

It is therefore unsurprising that many of Frege’s best known doctrines involve the notion of truth: Truth-values are the references of sentences; concepts are functions from objects to truth-values; negation and the conditional are truth-functions; the sense of a sentence is its truth-condition.

These doctrines are, of course, inter-related and mutually supporting. The thesis that sentences refer to truth-values implies that the sentential connectives denote truth-functions, given that they denote some sort of function or other: The sentences that occur as arguments of the conditional refer to truth-values, and so does the conditional itself; so if the conditional denotes a function, then it must denote a function from truth-values to truth-values. Similarly, the thesis that predicates refer to functions from objects to truth-values follows from the thesis that

[12] Strawson’s terminology is thus a bit confusing, but he explicitly does *not* mean by a ‘statement’ an *utterance*: That truth applies to utterances is, Strawson (1950, §1) thinks, one of the central errors in the theory of truth due to J. L. Austin (1951) that he is criticizing.

[13] These remarks were written around 1918 and so are from very late in Frege’s career. Similar remarks are found, however, both in Frege’s unpublished “Logic” from 1897 (Log97, p. 128) and, a little less explicitly, in an earlier piece by the same title (Log79, p. 3).

sentences refer to truth-values, given three of Frege's other commitments: that predicates refer to functions; that names refer to objects; and that the semantic value of a sentence "Fa" is the result of applying the function to which "$F\xi$" refers to the object to which "a" refers.[14] For consider such a sentence. We know that '$F\xi$' must denote a function and that 'a' must denote an object. So the arguments to the function '$F\xi$' are objects, and if the reference of 'Fa' is to be a truth-value, then the values of the function must be truth-values. So '$F\xi$' denotes a function from objects to truth-values.[15]

We have argued elsewhere (Heck and May, 2010, §4) that the thesis that the sense of a sentence is its truth-condition is also a consequence of Frege's conception of how thoughts are composed from senses, given that sentences refer to truth-values. And it is, or so we suggested in the last section, because sentences refer to truth-values—a claim that is itself a reflection of the intimate link between judgement and truth—that logic is constitutively concerned with truth. The most fundamental of Frege's mature doctrines about truth is thus the thesis that sentences refer to their truth-values or, as we shall call it, the Truth-Value Thesis. It is on it that the other doctrines rest.

It is important to distinguish the Truth-Value Thesis, as Frege himself does not, from the related thesis that truth-values are *objects*. As Dummett (1981, ch. 12) has emphasized, given the fact that, for Frege, ontological categories supervene on syntactic ones, the latter thesis is all but equivalent to the claim that sentences are a sort of proper name. That is, Frege holds quite generally that what kind of entity something is—an object, concept, or function of whatever type—is determined by what kind of expression might refer to it (Dummett, 1981, chs. 3–4). If so, then the claim that truth-values are objects becomes the claim that sentences (which are the kinds of expressions that refer to truth-values) are of the same logical type as the expressions normally fit to refer to

[14] This last doctrine reflects Frege's view that semantic composition is, in a sense we try to explain elsewhere (Heck and May, 2013, §4), 'internal' to the semantics of predicates.

[15] Frege holds similar views about predicates of other logical types. For example, a two-place predicate like 'ξ loves η' refers, on Frege's view, to a two-place function from objects to truth-values: it refers to a two-place, first-level concept. Higher-level concepts take concepts (more generally, functions) as arguments and have truth-values as their values. Quite generally, then, predicates refer to functions from arguments of some appropriate type (or types) to truth-values.

objects, proper names. It should be obvious that this syntactic thesis is independent of the thesis that sentences refer to their truth-values.[16]

As Dummett also makes clear, the claim that sentences *refer* to things, while it sounds odd to the modern ear, should not be doubted on that ground.[17] Frege's notion of reference, though grounded intuitively in the relation between a name and its bearer, is a technical one, to be explicated, ultimately, in terms of the theoretical work it is supposed to do (Dummett, 1981, pp. 196–203). That work is semantic, which is why Dummett suggests we might use the term "semantic value" for the general notion Frege has in mind, perhaps reserving "reference" for the special case of proper names.

The role that truth is supposed to play in logic is clearest from Frege's own elaboration of his formal system in *Grundgesetze der Arithmetik*. Part I of that book, entitled "Exposition of the System", is intended, first and foremost, to introduce and explain the system of logic in which Frege proposes to give his "Proofs of the Basic Laws of Cardinal Number" in Part II. But Frege's goal in Part I is not, in our view, just expository.[18] Frege's overall goal in *Grundgesetze* is to show "that arithmetic is merely logic further developed" (Gg, v. I, p. vii), but the provision of formal proofs of axioms for arithmetic, though necessary for this goal—since, otherwise, something intuitive might intrude unnoticed (Gl, §90)—is not sufficient. The question remains whether the Basic Laws from which the proofs themselves proceed, and the inferences that constitute those proofs, are logical in character (PCN, opp. 362–3). Only if that is so will the theorems of Frege's system be guaranteed to be truths of logic.

The real goal of Part I, then, is to secure this additional claim: to demonstrate that each of Frege's six Basic Laws is true and to show that each of his rules of inference preserves truth. These arguments depend, as they must, upon a series of stipulations Frege makes regarding the meanings of the primitive expressions of his formal language, his 'begriffsschrift' (cf. Heck, 2012a, §2.2). For example, he stipulates that a

[16] In their semantics textbook *Knowledge of Meaning*, Larson and Segal (1995) treat sentences as referring to truth-values, but do not regard sentences as being a special sort of proper name.

[17] Frege's German term is "*bedeuten*", whose most natural translation is "meaning". But that does not help a great deal, since the thesis that sentences 'mean' their truth-values sounds odd for a different reason.

[18] Heck (2012a, Part I) gives detailed arguments for this claim.

conditional

$$\begin{array}{l}\ulcorner\!\!-\!\!-\,\Gamma\\ \llcorner\!\!-\,\Delta\end{array}$$

is to have as its value the False only if Δ is the True and Γ is the False (Gg, v. I, §11), and he then uses this stipulation to show that his Basic Law I:

$$\vdash\!\!\begin{array}{l}\ulcorner\!\!\ulcorner\!\!-\,p\\ \;\;\llcorner\!\!-\,q\\ \llcorner\!\!-\!\!-\,p\end{array}$$

is true (Gg, v. I, §18) and that his first method of inference, *modus ponens*, preserves truth (Gg, v. I, §14).

Frege makes similar stipulations concerning the other primitives,[19] and these stipulations together comprise an informal, but nonetheless rigorous,[20] semantic theory for begriffsschrift. Frege goes on to argue in §31 that what he has stipulated concerning his primitives is adequate to assign a unique denotation to every expression of begriffsschrift. Specifically, his stipulations are adequate to assign a unique truth-value to every sentence and, therefore, to determine the condition under which any given sentence of the language will be true (Gg, v. I, §32). In that sense, Frege's semantics is 'materially adequate', in roughly Tarski's sense: For each sentence S of the language, it allows us to prove something of the form

$$S \text{ refers to the True iff } p$$

where p is a translation of the formal sentence S into our informal meta-language.

To be sure, the argument Frege gives in §31 is fatally flawed. If, indeed, every sentence of begriffsschrift had a unique truth-value as its reference, then we could argue for the consistency of the begriffsschrift as follows: It follows from the stipulations Frege makes concerning his primitives that all of his Basic Laws refer to the True and, moreover, that his rules of inference preserve the property of referring to the True; but then, by a simple induction, every theorem refers to the True; yet

[19] That concerning the horizontal is in v. I, §5; negation, §6; identity, §7; the first-order universal quantifer, §8; the smooth breathing, §9; the definite article, §11; and the second-order universal quantifer, §24.

[20] One almost as rigorous as the semantics for the calculus of classes that Tarski develops in §3 of "The Concept of Truth in Formalized Languages" (Tarski, 1956), which is also stated in an informal meta-theory and is in no sense 'formal'.

there is at least one sentence (e.g., $\overset{\mathfrak{a}}{\smile}\ \mathfrak{a} \neq \mathfrak{a}$) that does not refer to the True; hence, not every sentence can be a theorem, and the begriffsschrift must be consistent. Since it isn't, something has gone wrong, a point Frege recognizes himself in one of his letters to Russell (PMC, p. 132).

But the fact that the argument of §31 is flawed does not imply that it does not tell us something important about how Frege thought about his formal language and its relation to truth. Indeed, although the argument is flawed, its shortcomings are due entirely to peculiarities connected with the stipulation Frege makes concerning the *spiritus lenis*, or 'smooth breathing', which is used to form names of value-ranges and which is the subject of the infamous Basic Law V. If we omit the smooth breathing, then Frege's argument *really does show* that every expression of the 'logical fragment' of begriffsschrift has been provided with a unique denotation and a definite truth-condition.[21]

The claim that sentences 'refer' to their truth-values is of course at the foundation of Frege's semantics for begriffsschrift: The 'semantic value' of a sentence is its truth-value, as Frege's stipulation concerning the conditional makes clear. In that respect, then, the Truth-Value Thesis is actually quite familiar: The sort of semantics we teach in elementary logic embodies the same idea, that the fundamental semantic fact about a sentence is that it is true or false, that is, that it has whatever truth-value it has. Nonetheless, the Truth-Value Thesis seems to many just strange, and it does not help that Frege gives almost no explicit argument for it. The few arguments he does give are in "On Sense and Reference", and they are weak.

The first of these arguments is that it only matters to us whether a name refers in so far as we are concerned with the truth-value of some sentence in which it occurs. That may be, but we are hardly "driven into accepting the truth-value of a sentence as constituting what it refers to" (SM, op. 34) by this sort of consideration. For it might well be that (i) we are only interested in the reference of a sentence when we are interested in its truth-value and (ii) what the reference of the sentence is depends upon the references of the names that are contained in it. If so, then we would indeed be interested in the reference of a name only when concerned with the truth-value of some sentence containing it. But if a sentence referred to a state of affairs, then (ii) would certainly be true,

[21] And one can then prove that the logical fragment of the begriffsschrift is consistent, in much the same way that Tarski (1956, p. 199, Theorem 7) proves that the calculus of classes is. This is not a terribly interesting result, however, since the logical fragment has a two-element model.

and one can at least imagine a plausible argument for (i), i.e., that we are interested in what state of affairs a sentence refers to only when we are interested in its truth-value. But if even if that were not so, then that would be because there were cases in which we were interested in the reference of a name because we were interested in what state of affairs a sentence 'expressed', even though we were *not* interested in whether that sentence was true. And then Frege's argument has a false premise.

Frege's second argument is a little better, but not much:

> If our supposition that the reference of a sentence is its truth-value is correct, then the latter must remain unchanged when a part of the sentence is replaced by an expression with the same reference. And this is in fact the case. . . . If we are dealing with sentences for which the meaning of their component parts is at all relevant, then what feature except the truth-value can be found that belongs to such sentences quite generally and remains unchanged by substitutions of the kind just mentioned? (SM, op. 35)

This is usually interpreted as an attempt to derive the Truth-Value Thesis from the compositionality of reference, that is, from the claim that the reference of a complex expression is determined by the references of its parts. But it too is unconvincing. What the argument certainly does show is that it is *consistent* with compositionality to take sentences to refer to their truth-values, but it does not follow that sentences refer to truth-values unless there is no other view with that is consistent with compositionality. And the view that sentences refer to states of affairs might well be thought to be just such a view. That they refer to Russellian propositions (or to sets of metaphysically possible worlds) is certainly such a view.

Why are Frege's arguments for the Truth-Value Thesis so terrible? The answer is simple: He doesn't really have a direct argument for it. But these are not the sorts of considerations that might plausibly have led him to it, anyway. A better indication of why Frege adopted the Truth-Value Thesis is a remark he makes in the Foreword to *Grundgesetze*: "Only a thorough engagement with the present work can teach how much simpler and more precise everything is made by the introduction of truth-values" (Gg, v. I, p. x). Frege thinks that the Truth-Value Thesis solves a lot of problems, and he thinks that it solves them better than anything else on offer. That is the real reason he adopts it.

To understand why Frege embraces the Truth-Value Thesis, then, we need to discover what problems he thinks it solves. And the key to that task is the realization that *not one* of the distinctive doctrines about truth that we listed at the beginning of this section is present in Frege's earliest work: Frege does not hold, in *Begriffsschrift*, that sentences refer to truth-values, nor that concepts are functions from objects to truth-values, nor that negation and the conditional are truth-functions. Indeed, there is hardly any mention of truth in Frege's writings before 1890.

In the sections that follow, then, we shall explore these various doctrines in an effort to understand how, and in response to what pressures, they emerge.

3 Concepts as Functions From Objects to Truth-Values

Let us focus first on Frege's doctrine that (first-level) concepts are functions from objects to truth-values. The idea that concepts are functions is already present in *Begriffsschift*. Frege presents the distinction between function and argument as the key to his new analysis of generality, and the way he explains that distinction (Bg, §9) is strikingly reminiscent of how he explains the distinction between function and object in his mature work (FC, op. 6). And, although Frege does not use the term "concept" in *Begriffsschrift*, when he introduces the distinction between function and argument, he does so by applying it to the sentence "Hydrogen is lighter than carbon dioxide". Frege says that, if we regard "hydrogen" as replaceable by other expressions, then "...'hydrogen' is the argument and 'being lighter than carbon dioxide' is the function..." (Bg, §9). Obviously, "being lighter than carbon dioxide" is a predicate, and so the upshot is that we are to regard predicates, logically, as being functions.

Frege thus regarded a simple sentence like "5 is prime" as being analyzable into a function and an argument. This might sound familiar, but it is important to appreciate that this distinction is actually very different from Frege's mature distinction between concept and object. Whereas the distinction between concept and object "is not made arbitrarily, but founded deep in the nature of things" (FC, op. 31), the distinction between function and argument "has nothing to do with the conceptual content [but] comes about only because we view the expression [of that conceptual content] in a particular way" (Bg, §9). What, on one way of viewing "5 is prime", we regard as a function, on another way of viewing it, we may regard as an argument (Bg, §10). If we imagine "5" replaced by other expressions, then we are regarding it as the argument. But we may also imagine "is prime" replaced by other expressions. Then we would be regarding it as the argument.

Frege could have made an analogous point in his mature period. He would then have regarded "5 is prime" as most fundamentally composed of a name, "5", and a concept-expression, "ξ is prime", where "ξ" indicates the 'incompleteness' that Frege then understood such expressions to have. But one can also regard the sentence as saying something like: Prime is something 5 is. In that case, one is thinking of the sentence as being composed of the first-level predicate "ξ is prime" and the second-level predicate "$5_x(\Phi x)$", a predicate that means something like: is instantiated by 5. But this is not how Frege is thinking of matters in *Begriffsschrift*. The analysis just described depends crucially upon the distinction between first- and second-level predicates and, more generally, upon the notion of incompleteness or unsaturatedness, neither of which is present in *Begriffsschrift*.

Perhaps what is most striking about Frege's early distinction between function and argument is that it is essentially a linguistic one. Frege's official view in *Begriffsschrift* would appear to be that functions are *expressions*. If, in the sentence "5 is prime", we take the argument to be "5", then the function is "the part that remains invariant *in the expression*" when we replace "5" by other names, such as "7" or "12" (Bg, §9, our emphasis). That is why we said, a few paragraphs ago, that the *predicate* is to be regarded as a function.[22]

From this point of view, the question what the values of the function "is prime" are supposed to be need never arise, and Frege never addresses himself to this question. The only relevant remark would seem to be in his introduction of his notation for generality, where he says that "$\vdash\!\overset{\mathfrak{a}}{\smile}\!\!-\, \Phi(\mathfrak{a})$" "stands for the judgement that, whatever we may take for its argument, the function is a fact" (Bg, §11). But it is utterly unclear what this is supposed to mean (and the problem does not seem to be with the translation). The *function* is a fact? Surely not. Frege must mean that, for each argument, the *value* of the function is a fact. But what is its value if it is not a fact? A non-fact? It would appear that Frege simply had not thought such matters through.

[22] For defense and elaboration of the claims made in the last few paragraphs, see our paper "The Function is Unsaturated" (Heck and May, 2013), especially §28.2.

Within just a couple years, however, Frege abandons the linguistic view of functions and replaces it with a more modern view that regards functions as mappings from arguments to values (Heck and May, 2013, §28.3). And once this new view is in place, Frege can no longer avoid the question what might be the values and arguments of the functions associated with predicates. Moreover, Frege no longer considers the argument of the function to be a name, but rather what the name stands for. And Frege held already in *Begriffsschrift* that the content of a proper name is the object it denotes (Bg, §8), so the arguments to what we shall call 'concept-functions' are objects, as one would expect.

What, however, might the values of concept-functions be? It is easy enough to deduce Frege's original answer by applying the sort of reasoning we used above. Consider a simple sentence "Fa". Applying the concept-function that is the content of "F" to the object that is the content of "a" should yield the content of the complex expression that they together constitute. But the content of a proper name has as its content an object; the content of a sentence "Fa" is a content that can become a judgement (Bg, §2), for short, a 'judgeable content'; so the function that is the content of "F" must take objects arguments and have judgeable contents as values. That is, Frege's early view must have been that concept-functions are functions from objects to judgeable contents.[23]

This view is strikingly reminiscent of Russell's notion of a propositional function (Russell, 1903, ch. VII) and is really quite elegant. Unfortunately, as now well-known considerations show, it has consequences that would have been unacceptable to Frege. If the content of

(1) The Evening Star is a planet

is the result of applying the function that is the content of "ξ is a planet" to the object that is the content of "the Evening Star", then its content must be the same as that of

(2) The Morning Star is a planet.

[23] Michael Beaney (2007) comes to the same conclusion. Additional evidence is provided by Frege's remark in the preface to *Grundgesetze* that conceptual content "now splits for me into what I call thought and what I call truth-value [as] a consequence of the distinction between the sense and reference of a sign" (Gg, v. I, p. x). Frege's language here obviously suggests that the roles of (i) sentential content and (ii) the value of a concept-function had previously been played by a single notion, and that this notion is what has now split (functionally, not mereologically) into sense and reference. We have discussed this point further elsewhere (Heck and May, 2006, §4).

But (1) and (2) cannot have the same content. The point is not simply that we have some 'intuition' that they do not (cf. Heck, 2014). The notion of content that is in play here is not an 'intuitive' one but one whose natural home is logic:

> . . . [T]he contents of two judgements may differ in two ways: either the consequences derivable from the first, when it is combined with certain other judgements, always follow from the second, when it is combined with these same judgements, or this is not the case. The two propositions, "The Greeks defeated the Persians at Plataea" and "The Persians were defeated by the Greeks at Plataea" differ in the first way. . . . Now I call that part of the content that is the *same* in both the *conceptual content*. . . . Everything necessary for a correct inference is expressed in full, but what is not necessary is generally not indicated; *nothing is left to guesswork*. (Bg, §3, emphasis in original)

Some readers, such as Robert Brandom (1994, pp. 94ff) and Michael Kremer (2010), have claimed to find in this passage an argument for the claim that, if A and B have the same consequences, then they must also have the same content. We don't read the passage that way, but the issue is not relevant at the moment.[24] What matters for our purposes is only that a weaker thesis clearly is present: If A and B have the same content, then they must have the same logical properties. If so, then since it is obvious that (1) and (2) have different logical consequences,[25] it follows that (1) and (2) must have different conceptual contents. We'll call this the 'substitution argument'.

How can Frege avoid the conclusion that (1) and (2) must have the same content? The conclusion depends upon just four assumptions:

(i) The content of a proper name is its bearer.

(ii) The contents of predicates are functions.

(iii) The content of a simple sentence "Fa" is the result of applying the concept-function that is the content of "F" to the content of "a".

[24] Even if we did read the passage that way, we would counsel against assuming that it had any implications whatsoever for Frege's mature views.

[25] We take this to be obvious ourselves—although it can be and has been denied (Marcus, 1995)—but, more importantly, we take it to have been obvious to Frege, i.e., we take it that Frege would have thought it obvious.

(iv) Logical properties are determined by content, so that sameness of content implies sameness of logical properties.

Only (iii) is dispensible: (ii) is the key to Frege's understanding of logical generality; (iv) is integral to Frege's conception of logic and its relation to content (May, 2006);[26] and (i) is central to Frege's understanding of identity as objectual. This last view is not present in *Begriffsschrift*, but it is in place by 1881 and is fundamental to Frege's logicism (May, 2001; Heck and May, 2006; Heck, 2019). So, as said, (iii) is what Frege must abandon: The content of "Fa" *cannot* be the result of applying the concept-function that is the content of "F" to the object that is the content of "a".[27] To put it differently: The values of concept-functions *cannot* be judgeable contents. So what are they?

We know, of course, what Frege's answer to this question will ultimately be: The values of concept-functions are truth-values, so that concepts are functions from objects to truth-values. So this thesis—and, with it, the Truth-Value Thesis—is partly a result of Frege's being forced to re-think the question what the values of concept-functions are in light of the substitution argument, which shows that they cannot be judgeable contents.

Another source of Frege's mature view lies in how he proposes to distinguish concepts from objects. As we mentioned earlier, the notion of unsaturatedness, or incompleteness, does not appear in *Begriffsschrift* or in any of Frege's (extant) writings before 1881. It clearly is present, however, by 1882, when Frege writes, in a letter to the philosopher Anton Marty: "A concept is unsaturated in that it requires something to fall under it; hence it cannot exist on its own" (PMC, p. 101). More important for our purposes, however, is an earlier remark that the one just quoted elaborates, that "it [is] essential for a concept that the question whether something falls under it has a sense" (PMC, p. 101), a question that Frege says would be senseless in the case of an object. What this suggests is

[26] It is (iv), of course, that Russellian responses to Frege's puzzle about substitution encourage us to abandon (Fine, 2007; Heck, 2012b).

[27] So long, that is, as Frege is working with an undifferentiated notion of content. As we shall see in Section 6, however, other problems lurk here, ones that will eventually force Frege to split the notion of content into sense and reference (see note 23). Then (i), (ii), and (iii) will hold when "content" means *reference*, but (iv) will hold only when "content" means *sense*. Whether (ii) and (iii) also hold when "content" means *sense* is controversial, though our view is that they do not (Heck and May, 2010, esp. §4). That (i) does not hold when "content" means *sense* is, of course, the very point of the distinction between sense and reference.

that Frege views concepts as sorting objects into two baskets: those that fall under the concept and those that do not. The view that concepts are functions from objects to truth-values fits naturally with this suggestion.

But such considerations cannot, on their own, have driven Frege to the view that truth-values are the values of concept-functions, because this view has deeply counter-intuitive consequences. In particular, it implies that concepts are extensional, so that there can be only one concept true of a given collection of objects. Suppose, for example, that (as philosophers' lore has it) the animals that are supposed to have kidneys are the same as the animals that are supposed to have hearts. Then the function that maps an animal to the truth-value True just in case that animal is supposed to have a kidney is the very same function as the one that maps an animal to the True just in case that animal is supposed to have a heart. So, if concepts are functions from objects to truth-values, then the concepts *renate* and *cordate* are the same concept, which is a deeply counterintuitive result. That is Frege's *mature* view, to be sure: concepts are extensional. But he did not start out with that view.

The view that concepts are *in*tensional is an almost immediate consequence of Frege's original view that concepts are functions from objects to judgeable contents. On that view, *is a renate* maps Joe to *Joe is a renate*, and *is a cordate* maps Joe to *Joe is a cordate*. Since the content *Joe is a cordate* has different logical properties from the content *Joe is a renate*, these are different contents, so the functions *is a cordate* and *is a renate* are distinct: They take different values for Joe as argument. Indeed, they take different values for *every* argument.

There is also direct textual evidence that Frege originally thought that concepts were intensional. In "Boole's Logical Calculus and the Begriffsschrift",[28] Frege implicitly uses the thesis that concepts are intensional as a premise in one of his main arguments:

> The concept 'planet whose distance from the sun lies between that of Venus and that of Mars' is still something different from the individual object the Earth, even though [the Earth] alone falls under the concept. Otherwise, you couldn't form concepts with different contents whose extensions were all limited to this one thing, the Earth. (BLC, p. 18)

[28] This paper was not published by Frege during his lifetime, but the editors of the *Nachlass* note that he submitted it for publication three times. We can therefore take it to express his considered views at that time.

The argument Frege is giving here intended to establish a conclusion that matters deeply to him: that we must distinguish concepts from objects. The argument for that conclusion is simple: If every concept true just of the Earth was identical with the Earth, then all concepts true just of the Earth would be the same; but they aren't—that is the intensionality premise—so concepts true just of the Earth aren't identical with the Earth.[29] Frege's implicit claim that you *can* "form concepts with different contents whose extensions [are] all limited to this one thing" is thus one on which he is willing to put quite a lot of weight. And yet, he feels no need to give an argument for it.

That the Boolean logicians treat objects as if they were concepts is, in fact, one of Frege's main criticisms of them. For Boole and his followers, sentences were constructed from predicates using the traditional forms of judgment. A universal affirmative judgment, for example, would be constructed by inserting predicates into the form: All ... are .... It might seem as if this would make the famous argument

All humans are mortal.
Socrates is a human.
Therefore, Socrates is mortal.

difficult to handle. But the Booleans regarded proper names as simply a special sort of predicate, ones that happen to be true of a single thing. In particular, "Socrates" is a predicate, one that is true just of Socrates.[30] The famous argument just mentioned could therefore be represented as:

All H are M.
All S are H.
Therefore, all S are M.

and its correctness is then evident.[31]

Frege repeats this criticism in *Die Grundlagen*, directing it at Ernst Schröder (a prominent German member of the Boolean school),[32] emphasizing that "... a concept does not cease to be a concept simply because

[29] As Richard Cartwright once pointed out in a lecture, though in a slightly different context, it is consistent with this argument that *one* concept true of just the Earth should be identical with the Earth. It's a nice question what additional premise is needed here. Perhaps, Cartwright suggested, it is some form of the Principle of Sufficient Reason.

[30] One might compare this idea to W. V. O. Quine's famous suggestion that we might verb all the names, e.g., replace "Pegasus" by "Pegasizes" (Quine, 1953).

[31] We have discussed this objection of Frege's in more detail elsewhere (Heck and May, 2006, §2).

[32] In 1877, Schröder had published a short exposition of the basic ideas behind Boolean logic (Schröder, 1877). Frege cites this book in at least two places (BLC, pp. 38–9; BLF,

only one single thing falls under it. . . ” (Gl, §51). And, of course, the principle that concepts must, quite generally, be distinguished from objects is one of the “three fundamental principles” that Frege lists in the introduction to *Die Grundlagen* as crucial to the argument of the book (Gl, p. x). The argument for this principle that Frege gives in “Boole’s Logical Calculus” does *not* appear in *Die Grundlagen*, however, and the reason is that Frege has by then abandoned the view that concepts are intensional. In a footnote attached to his definition of number, Frege writes:

> I believe that for “extension of the concept” we could write simply “concept”. But this would be open to two objections:
>
> 1. that this contradicts my earlier statement that the individual numbers are objects. . . ;
> 2. that concepts can have identical extensions without themselves coinciding.
>
> I am, as it happens, convinced that both these objections can be met; but to do this would take us too far afield for present purposes. (Gl, §68, fn. 1)

It is extremely unfortunate that Frege does not say how he would have answered the second of these objections, that is, how he now proposes to defend the view that concepts are extensional. But the prior question is why he has abandoned his earlier view that concepts are intensional. The later view is far less plausible, so Frege must have had very good reason to commit himself to the claim that concepts are extensional. What sort of change in his thinking might have forced him to do so? We can see no answer other than: Adoption of the view that concepts are functions from objects to truth-values. That would imply that concepts with the same extension must ‘coincide’ and so would provide Frege with the resources to answer the second objection. So the thesis that concepts are functions from objects to truth-values must therefore be in place by 1884, at the latest.

We pause to emphasize this point, because it is extremely important, and much of what we shall have to say below depends upon it. So here

p. 49). As we’ll note below, Frege does not seem to have known of Boole’s work when he wrote *Begriffsschrift*. He seems to have learned of it from reviews published, by among others, Schröder (1972) himself.

it is again: The view that concepts are functions from objects to truth-values, and so one core piece of the Truth-Value Thesis, is already in place by *Die Grundlagen*. Our argument for this claim is admittedly indirect, but we simply cannot see how else one might explain why Frege would have abandoned the natural view that concepts are intensional in favor of the deeply problematic view that concepts are extensional, nor how else one might suppose that Frege would have proposed to defend his new view. The problem is made all the more difficult by the fact that so much else has to be held fixed, e.g., the view that concepts are functions, due to the role that claim plays, throughout Frege's career, in his analysis of generality.

There are a couple of counter-suggestions that might be made. As Frege indicates in "On Concept and Object" (CO, op. 195, n. 6), his response to the first objection is that the phrase "the concept F", which is what one finds in the text on which Frege is commenting, refers not to a concept but to an object (perhaps to the extension of that concept). One might therefore suggest that Frege's answer to the second objection would be similar: He does not think that "concepts can have identical extensions without themselves coinciding"; he just thinks that *the* concept F cannot differ from *the* concept G unless they have different extensions, since these phrases themselves refer to extensions, not to concepts. But we do not find that reading very plausible. This interpretation fails to explain why the argument from "Boole's Logical Calculus" does not reappear in *Die Grundlagen*. Moreover, Frege seems to be indicating that he thinks it is *false* that "concepts can have identical extensions without themselves coinciding", not that the second objection fails to engage with what he has said.

Alternatively, one might note that the view that concepts are functions from objects to truth-values does not, by itself, entail that concepts are extensional: Functions themselves might be intensional. If so, then concepts will be intensional, too. So the natural suggestion is that Frege had regarded *functions* as intensional before *Die Grundlagen* but now regards them as extensional, and that is why he takes concepts to be extensional in *Die Grundlagen*. The problem, however, is that the extensionality of functions does not, by itself, entail the extensionality of concepts. The argument rehearsed earlier for the intensionality of concepts did not depend upon the claim that functions are intensional. If concepts are functions from objects to judgeable contents, then *is a cordate* maps Joe to *Joe is a cordate* and *is a renate* maps Joe to *Joe is a renate*, and so those functions map Joe to different values; it is irrelevant whether the functions themselves are extensional. What is right, though, is that the extensionality of concepts follows only from the combination of these two views: that concepts are functions from objects to truth-values *and* that functions are extensional. Frege's abandonment of the intensional view of concepts makes sense, therefore, only if he has committed himself to *both* of these views by *Die Grundlagen*. It is the commitment to the former view that matters here, though we do indeed take these same considerations to show that the latter view was also in place by 1884.

For what it's worth, we think this probably does represent another change in Frege's views. As mentioned earlier, Frege does not clearly distinguish functions from the expressions that denote them in *Begriffsschrift*. If so, then it would have been natural for him to regard $x^2 - 1$ and $(x+1)(x-1)$ as different functions. Frege abandons this view, presumably, for much the same reason he abadons the view that identity is a relation between names: his emerging opposition to formalism (May, 2001), which is firmly in place by *Die Grundlagen* (Gl, §§92–99) and which is the entire focus of his paper "Formal Theories of Arithmetic", which was published a year later. And once Frege has distinguished functions from expressions, it becomes very natural for him also to regard functions as extensional. As he himself points out (Gg, v. II. §147), mathematicians of his day, no less than of ours, often said such things as that $x^2 - 1$ and $(x + 1)(x - 1)$ are the same function. Frege objected to such language on the ground that identity is a relation between objects, and functions are not objects. But the functions may be regarded as standing in a relation appropriate to their type that "corresponds to" identity (CSM, p. 121), and that relation treats functions as extensional.

So the view that concepts are (extensional) functions from objects to truth-values seems to have been in place by 1884. The fact that this view has such counter-intuitive consequences makes it a nice question, again, why Frege came to hold it. The answer, or so we have suggested, is that Frege had by then arrived at the view that concepts are functions from objects to truth-values. The question then becomes why Frege might have adopted that view. And the answer, or so we have suggested, is that it is a consequence, for reasons we have already considered, of the Truth-Value Thesis: If (i) "Fa" denotes a truth-value, (ii) "a" denotes an object, and (iii) the incomplete predicate symbol "$F\xi$" denotes a function from the denotation of its argument to the denotation of the result of completing it with that argument, then it simply follows that concepts—what incomplete predicate symbols denote—are functions from objects to truth-values. Indeed, the claim that concepts are functions from objects to truth-values is simply equivalent to the Truth-Value Thesis, given Frege's other commitments.

We are suggesting, therefore, that Frege had settled upon the Truth-Value Thesis by 1884. The question we must explore next, then, is why that might be.

4 An Antinomy in *Begriffsschrift*?

We argued earlier that Frege was forced to abandon his original view that concepts are functions from objects to judgeable contents by the substitution argument, which purports to show that these two sentences:

(1) The Evening Star is a planet.

(2) The Morning Star is a planet.

must have the same content. A version of this argument, it turns out, can be presented in the formal theory of *Begriffsschrift*. What makes the formalization possible is the fact that there is, in the formal language of *Begriffsschrift*, a symbol "$\equiv$" for what Frege calls 'identity of content': The official reading of "$A \equiv B$" is "the sign A and the sign B have the same conceptual content, so that we can everywhere put B for A and conversely" (Bg, §8, emphasis removed). But the substitution argument involves little more than an applicaton of Leibniz's Law, which is proposition (52) of *Begriffsschrift*. The argument thus turns out to be very short indeed:[33]

$$\begin{aligned}
&\vdash c \equiv d \to [(Fc \equiv Fc) \to (Fc \equiv Fd)] \\
&\vdash Fc \equiv Fc \\
&\vdash c \equiv d \to Fc \equiv Fd
\end{aligned}$$

The first formula is the instance of Leibniz's Law for "$Fc \equiv F\xi$"; the second is an instance of proposition (54) of *Begriffsschrift*; the conclusion then follows by *modus ponens*, more or less.[34]

It is hard to imagine that Frege would not have discovered this antinomy himself. For one thing, this sort of substitution argument was dear

[33] We shall use modern notation here, and in other places where doing so makes no significant difference.

[34] As a special case—take "$F\xi$" to be "$c \equiv \xi$"—we have:

$$c \equiv d \to [(c \equiv c) \equiv (c \equiv d)]$$

to Frege's heart, being at the very center of "On Sense and Reference". For another, arguments of similar structure occur in *Begriffschrift* itself. The proof of symmetry, which is proposition (55), involves the same sequence of steps, except that the instance of (52) that we need is the one for "$\xi \equiv c$":

$$\vdash c \equiv d \to (c \equiv c \to d \equiv c)$$
$$\vdash c \equiv c$$
$$\vdash c \equiv d \to d \equiv c$$

Moreover, the conclusion of the argument we sketched, $c \equiv d \to Fc \equiv Fd$, is relatively obvious, since it is just expresses a principle of compositionality for content: If you replace a part of a sentence with another with the same content, then the content of the whole will be unchanged.

Indeed, one might have thought that Frege should have formulated Leibniz's Law not as:

$$\vdash c \equiv d \to (Fc \to Fd), \tag{52}$$

but instead as:

$$\vdash c \equiv d \to Fc \equiv Fd, \tag{52$'$}$$

which is simply the conclusion of the argument we are discussing. He cannot do so, however, because we would not then know how to 'do anything' with the consequent: The problem, in effect, is that we would not then have an elimination rule for "$\equiv$". It is (52) that functions as such an elimination rule. But (52′) remains the natural form of Leibniz's Law.

How can Frege respond to this antimony? How, that is, can he prevent (52′) from being a theorem of the system? His options are limited. One idea would be to strengthen what "$c \equiv d$" means, taking it to be true only if "c" and "d" have the same content in some sense stronger than their simply denoting the same object. But even in *Begriffsschrift*, Frege clearly thinks he needs to be able to express extensional identity—that would seem to be the point of the discussion of identity in §8—and Frege's emerging logicism will quickly require a notion of objectual identity, anyway (May, 2001). Even if that were not so, however, simply

which amounts to a formal refutation of the theory of identity in *Begriffsschrift* §8: Any true identity-statement will have the same content as a triviality. Kremer (2010) also notes the derivability of this formula. His proof involves substituting $(c \equiv c) \equiv (c \equiv \xi)$ into (52). See also May (2012).

weakening "$c \equiv d$" to "$c = d$" will not help if "$=$" it is still subject to the same laws as "$\equiv$", since then the antinomy will simply resurface as:

$$\vdash c = d \rightarrow Fc \equiv Fd$$

So Frege has to do something about "$\equiv$" as it occurs in the consequent, that is, as it appears between sentences.

One option that will seem natural to contemporary readers is to declare that "$\equiv$" creates an opaque context and so that "$Fc \equiv F\xi$" is not a valid substituend in Leibniz's Law. And, in some sense, Frege has no choice but to pursue this option. So long as "...has the same content as..." can legitimately be substituted into Leibniz's Law, the antimony will remain. So we might see Frege's concern with intensional contexts in the later parts of "On Sense and Reference" as part of his response to the antinomy: It amounts to an explanation of why intensional contexts, in general, cannot legimately be substituted into Leibniz's Law.

Still, Frege never explores anything like intensional logic, so the suggestion we are considering must ultimately be that Frege should simply remove "$\equiv$" from his system, or else to re-explain it so that it does not create an opaque context. As we have said, this is undoubtedly something Frege needs to do. But again, in *Begriffsschrift*, Frege seems to think that he needs some such relation as that expressed by "$\equiv$" when it appears between sentences: It plays a crucial role in the statements of definitions, which are supposed to establish an identity of content. For Frege, then, it remains an important question how "$\equiv$" should be weakened: How, in the case of sentences, can we allow "$A \equiv B$" to be true, even though "A" and "B" do not have the same content?

We might try re-interpreting "$\equiv$" in terms of mutual implication: either defining "$A \equiv B$" as "$(A \rightarrow B) \wedge (B \rightarrow A)$" or else establishing axioms that would have the same effect.[35] That would certainly have resolved the antinomy, but, so far as we know, Frege never considers this option, even though it would have served his immediate purposes in *Begriffsschrift*.[36] As Frege mentions in "Boole's Logical Calculus", he uses the triple bar "between contents of possible judgement almost exclusively to stipulate the sense of a new designation" (BLC, pp. 35–6).

[35] As Alessandro Bandeira Duarte has pointed out, there is no way to derive the latter from the former in *Begriffsschrift*. To see this, note that all the axioms remain true if "$\equiv$" is taken to express some stronger notion of equivalence.

[36] Though his concerns are very different, this kind of suggestion is made by Hugh MacColl (1880, p. 51). We do not have any reason to think Frege read this paper, however, which is not the one that Schröder cites in his review of *Begriffsschrift* (Schröder, 1972).

And, in *Begriffsschrift*, Frege only ever uses a definition $\Vdash A \equiv B$ to infer the conditionals $\vdash A \rightarrow B$ and $\vdash B \rightarrow A$. However, re-interpreting "$\equiv$" in terms of mutual implication would leave us with no obvious reason to accept the sentential form (as it were) of proposition (52),[37] which is Frege's form of Leibniz's Law. Granted, as concerns sentential variables, (52) is not used essentially in *Begriffsschrift*, except to permit inferences from a definition to the corresponding conditionals. Given the definition $\Vdash A \equiv B$, Frege then cites

$$\vdash A \equiv B \rightarrow (A \rightarrow B),$$

which is a kind of degenerate instance of (52), with "$F\xi$" being replaced simply by "ξ". This inference is made, for example, in the derivations of propositions (75), (89), and (105); a related inference is involved in the proofs of (68) and (100).

To the modern eye, such an application of Leibniz's Law just looks weird. But, for Frege, (52) is what justifies the arbitrary replacement of the definiendum by the definiens, no matter where or how deeply nested the former might be. For example, Frege would have been perfectly happy to use (52) to derive a proposition like:

$$\vdash A \equiv B \rightarrow [(C \rightarrow A) \rightarrow (C \rightarrow B)],$$

where now "$F\xi$" has been replaced by: $C \rightarrow \xi$. Though such inferences do not appear in *Begriffsschrift*, the further development of Frege's system can be expected to require them, and there are lots of inferences of this sort in *Grundgesetze*.[38] Such a use of (52) is actually dispensible, of course, but that it is always dispensible is a significant meta-theorem, and it is certainly not the sort of thing Frege anywhere proves.[39]

Re-interpreting "$\equiv$" in terms of mutual implication would also abandon something else Frege might have thought important: the idea that

[37] This might nowadays be formulated as: $\vdash (A \rightarrow B \wedge B \rightarrow A) \rightarrow (\ldots A \ldots \rightarrow \ldots B \ldots)$, where the dots indicate some context in which A and B occur, and in which the latter has replaced the former.

[38] More precisely, Frege uses the equivalent of (52), his Basic Law III, to justify the arbitrary replacement of an expression with one proven equivalent to it, where the equivalence often rests upon a definition. See, for example, theorems (22), (23), and (33), and how they are applied.

[39] In most settings, it is enough to have Leibniz's Law for atomic formulas. One can then show, by induction on the complexity of expressions, that it will also be provable for complex formulas. Something similar is true in the sentential case. But, again, showing this kind of thing takes real work, and it is far from obvious. Indeed, because of the presence of "$\equiv$", it is not even clear that this would hold in *Begriffsschrift*. See note 35.

"$\equiv$" expresses a sort of identity. This element of Frege's view is not what is responsible for the antinomy. The antinomy emerges not from the formal result

$$\vdash c \equiv d \to Fc \equiv Fd$$

but from how the consequent is being interpreted, in particular, from the idea that its being true requires "Fc" and "Fd" to have the same content. So, again, resolving the antinomy means weakening that relation: It will have to be possible for "$A \equiv B$" to be true when A and B do not have the same content. But, if we are to preserve the idea that "$A \equiv B$" expresses some form of identity, then what is it that is the same?

That this is the right question to ask may well have been confirmed for Frege by the work of Boole and his followers, and Frege was intensely engaged with the work of the Booleans in the early 1880s. Boolean logic is an equational calculus, in which we find such results as[40]

$$A \times (B + C) = (A \times B) + (A \times C)$$

this being the Boolean version of the distributivity of conjuction over disjunction. Frege was generally quite critical of the Booleans' emphasis on analogies between arithmetic and logic (AimCN, pp. 93–4; BLC, pp. 13–5). But surely he would have noticed the similarity between the Booleans' use of "$=$" and his own use of "$\equiv$". It might even have seemed to him confirmation of his view that "$A \equiv B$" expresses some sort of identity. So the question again becomes: What is the same?

With this question in mind, consider again some remarks from "On Sense and Reference" that we quoted above:

> If our supposition that the reference of a sentence is its truth-value is correct, then the latter must remain unchanged when a part of the sentence is replaced by an expression with the same reference. And this is in fact the case. ...If we are dealing with sentences for which the meaning of their component parts is at all relevant, then what feature except the truth-value can be found that belongs to such sentences quite generally and remains unchanged by substitutions of the kind just mentioned? (SM, op. 35)

We earlier dismissed this as an unsuccessful attempt to derive the Truth-Value Thesis from the compositionality of reference. But we can now

[40] Notation varies among the various contributors to this tradition. We'll use the one in the text as it is relatively familiar.

see that there is more to it. Frege is gesturing in the direction of the antinomy we have been discussing and indicating how he thinks it ought to be solved. Unfortunately, this observation helps only a little, since, while Frege does point out that truth-values will do the job, he still gives us no reason to suppose that *only* truth-values will do the job.

It is still not clear, therefore, why Frege chooses truth-values to play this role. To answer that question, we need to look at a different aspect of the Truth-Value Thesis: one that is connected with the fact that sentences can occur as parts of other sentences.

5 Sentential Connectives as Functions

In their landmark study *The Development of Logic*, William and Martha Kneale (1962, pp. 420, 531) credit Frege with the discovery of truth-tables. They are in good company. Ludwig Wittgenstein too seems to attribute this discovery to Frege, though he is critical of Frege's understanding of truth-tables. He is reported to have said in 1935:[41]

> It is important to see that [the truth-table for disjunction] says nothing about the function $p \vee q$, but is another way of writing it. When Frege explained such functions by listing the truth-values of the arguments in columns on one side and the function on the other, it looked as though he had said something *about* the function. But instead he had defined it, given another notation for it. (Wittgenstein, 1979b, pp. 135ff)

The view Wittgenstein is attributing to Frege is of course very widespread today. The truth-tables, many of us would suppose, are bound up with the semantics of propositional logic; they embody a conception of how the semantic value of a complex formula depends upon, and is determined by, the semantic values of its parts. Wittgenstein, on the other hand, regarded truth-tables as a contribution to *syntax*. He would have regarded

p	q	
T	T	T
T	F	T
F	T	T
F	F	F

[41] Thanks to Michael Kremer for helping us locate this passage.

as being a formula—a 'propositional sign'—one we might otherwise write: $p \vee q$ (Wittgenstein, 1961, 4.442).

It is not our present purpose to try to resolve this disagreement, though it will be clear that we side with Frege. What we want to discuss here, rather, is the source of Frege's view that the sentential connectives—in particular, negation and the conditional—are truth-functions. And here again, the key to the investigation lies in the realization that this view is not present in Frege's earliest work.

Frege emphasizes in *Begriffsschrift* that, when we are considering a binary sentential connective, we must distinguish four possible cases, corresponding of course to the four lines of the truth-table. But although a table of sorts does appear in *Begriffsschrift*, *truth*-tables do not, since Frege does not mention the notion of truth in this connection: Frege simply does not say, in *Begriffsschrift*, that a conditional is false only when its antecdedent is true and its consequent is false. Frege's explanation of the conditional reads, rather, as follows:

> If A and B stand for contents that can become judgements. . . , there are the following four possibilities:
>
> (1) A is affirmed and B is affirmed;
> (2) A is affirmed and B is denied;
> (3) A is denied and B is affirmed;
> (4) A is denied and B is denied.
>
> Now
>
> $$\begin{array}{l} \vdash\!\!\!\top A \\ \;\;\,\llcorner B \end{array}$$
>
> stands for the judgement that *the third of these possibilities does not take place, but one of the other three does*. (Bg, §5, emphasis original; see also BLC, p. 35)

Frege does not use this sort of language in his explanation of negation in *Begriffsschrift*, but nor does he speak of truth:

> If a short vertical stroke is attached to the content stroke, then this will express the circumstance that *the content does not take place*. So, for example,
>
> $$\vdash\!\!\!\top A$$
>
> means "A does not take place". (Bg, §7)

The suggestion that negation and the conditional express truth-functions is therefore wholly absent from *Begriffsschrift* and, indeed, does not appear in any of Frege's extant writings from before the 1890s.

Truth-functions appear for the first time in the lecture *Function and Concept*. Shortly after completing the argument that functions are 'unsaturated', Frege suggests that we should think of concepts too as functions and then raises the question what their values might be, explaining that he takes them to be truth-values (FC, op. 13). Once the truth-values have been admitted as values of functions, however, it is then natural also to consider functions for which they are among the arguments. Interesting cases include negation and the conditional, and Frege clearly and explicitly explains these as truth-functions (FC, opp. 20ff).[42] That said, however, so far as we know, nowhere in his later writings does Frege give the sort of 'tabular' account that both Wittgenstein and the Kneales mention. Their crediting Frege with the discovery of the truth-tables thus appears to rest upon their conflating the tabular presentation in *Begriffsschrift* with the truth-functional explanation of the connectives in Frege's mature work. Putting these together doesn't require great insight, but the fact is that truth-tables as such do not appear in Frege's work. Moreover, there is no reason to suppose that Frege realized, as both Wittgenstein and Emil Post (1921) clearly did, that truth-tables can be used to decide the validity of an arbitrary propositional formula. There is therefore no basis for attributing the discovery of truth-tables to Frege. That honor belongs to Wittgenstein and Post, as is now widely appreciated. Still, the discovery of truth-*functions* is Frege's, as we shall now see.

[42] The fact that, in his mature period, Frege regards truth-values are objects complicates the situation, of course, since something like

$$\begin{array}{l}\vdash\!\!\top\; 1\\ \;\;\llcorner\; 0\end{array}$$

is not only well-formed but true. This is because the horizontal strokes that form part of the sign for the conditional are 'horizontals' in Frege's proprietary sense (Gg, v. I, §12), and so we may effectively regard the arguments as always being prefixed by horizontals and so as being truth-values. This fact plays an important role in the argument of §31, when Frege writes: "According to our specifications the names '$\top\,\Delta$' and '$\begin{array}{l}\top\,\Gamma\\ \llcorner\,\Delta\end{array}$' always have references if the names '$-\!\!-\,\Delta$' and '$-\!\!-\,\Gamma$' refer to something", thus disposing of the question whether negation and the conditional have a reference. Frege makes a similar move in the argument of §10.

Perhaps surprisingly, the idea that the sentential operators express functions is also absent from *Begriffsschrift*. Frege nowhere says that they do express functions, and that is not plausibly an oversight, given how concerned Frege is to establish the importance of the notion of function to logic. And there is indirect, but nonetheless impressive, evidence that Frege did not regard the connectives as expressing functions when he wrote *Begriffsschrift*.

One route to this conclusion would begin with a suggestion due to Øystein Linnebo (2003): that the Frege of *Begriffsschrift* shared with Immanuel Kant (from whom he presumably inherited it) the doctrine that logic is concerned only with the form of thought and not with its content. The notion of function would then belong only to content, whereas the conditional is part of the form, and so conditionals would not be treated as functions. The matter is unclear, however. The logical development in *Begriffsschrift* is, indeed, in some sense concerned only with form, but this may reflect a decision on Frege's part first to show what of mathematical significance follows strictly from matters of form before turning, as he does in *Die Grundlagen,* to the particulars of content.

There is, however, a different route to much the same conclusion. As Frege saw the matter in 1881, a "highly developed language" must contain two kinds of signs, "material" and "formal". In particular, a *lingua characterica* for mathematics must contain both sorts of elements. The material part of the language will include such symbols as "$+$" and "$\times$". Frege compares the formal part of the language to the "prefixes, suffixes, and auxiliary words" of natural language (BLC, p. 13), and the traditional view was precisely that such expressions had no meaning of their own. But the formal part, for Frege, includes most especially the logical symbols, so it appears that he regards them as being essentially without meaning and so certainly not as expressing functions. Rather, they would have been treated as syncategorematic.

Why, then, did Frege abandon this traditional view? Frege's emerging logicism would have given him one reason. Arithmetical truths are supposed to be logical truths, and Frege certainly did not think that arithmetic is without content. Indeed, in "Formal Theories of Arithmetic", Frege carefully explains the difference between two senses in which a theory of arithmetic can be 'formal', denying explicitly that arithmetic is formal in the traditional sense (FTA, opp. 97ff).

Another reason is to be found in Frege's reaction to criticisms of *Begrifsschrift* itself and, in particular, in his belated encounter with the work of Boole. Six reviews of *Begriffsschrift* were published,[43] among which was a condescending review written by Schröder (1972) in which he accused Frege of having ignored the accomplishments of Boole and his followers:

> With the exception of what is said... about 'function' and 'generality' and up to [Part III], the book is devoted to the establishment of a formula language that essentially coincides with Boole's mode of presenting *judgements* and Boole's calculus of judgements, and which certainly in no way achieves more. (Schröder, 1972, p. 221, emphasis original)

Schröder (1972, p. 220) speculates that Frege was ignorant of Boole's work, as does John Venn (1972, p. 234) in his review.[44] They were almost certainly right. As Tyrell Bynum (1972, pp. 77–8) points out, Frege took no courses in logic as a student, and some of his claims to originality reveal his ignorance of Boole.

It is no surprise that Frege should emphasize in his various replies (AimCN; BLC; BLF) how badly Schröder has failed to appreciate the significance of what Frege has to say "about 'function' and 'generality' ". But Frege's response is not limited to this point. His more potent criticism is that Boole's treatment of logic is fundamentally misconceived. For Boole, logic consists of two parts, which he calls the 'calculus of judgements' and the 'calculus of concepts'. The former is essentially what we now know as propositional logic; the latter has the same scope as traditional Aristotelian logic, treating of the sorts of relations between concepts expressed by "All F are G", "Some F are G", and so forth. Boole treats the calculus of concepts as fundamental and attempts to reduce the calculus of judgements to it. Frege argues not only that Boole's attempted reduction fails but, more generally, that Boole is wrong to treat the calculus of concepts as fundamental.[45]

Here is how Boole's reduction works. The languages of the two calculi are the same. Both contain expressions of the forms "$A \times B$",

[43] These are collected in the edition of *Begriffsschrift* edited by Tyrell Bynum (Frege, 1972).

[44] Venn never seems to have gotten over it. In the second edition of his *Symbolic Logic*, published in 1894, he writes of *Begriffsschrift*: "Here... we have an instance of an ingenious man working out a scheme—in this case a very cumbrous one—in apparent ignorance that anything better of the kind had ever been attempted before" (Venn, 1894, pp. 493–4). Venn seems to have known nothing of Frege's later work.

[45] Frege (Schr) would later make related criticisms of a somewhat different attempt that Schröder makes to found Boolean logic upon something akin to mereology.

"$A + B$", "$\bar{A}$", and so forth. In the calculus of concepts, the variables are supposed denote classes (or extensions of concepts):[46] subsets of the 'universe of discourse'. The operations are then interpreted set-theoretically, in the now familiar way: Multiplication is intersection; addition is union;[47] the bar represents the complement relative to the universe. Precisely how the operations were to be interpreted in the calculus of judgements was more controversial. One might expect that they would denote truth-values, but that would be wrong.[48] The whole point—what makes Boole's reduction possible—is that they again denote *classes*. What is controversial is what these classes contain as members, i.e., what sorts of things the universe of discourse comprises in the calculus of judgements. In *The Mathematical Analysis of Logic*, Boole (1847, pp. 48ff) takes the sentential variables to denote sets of 'cases' or 'circumstances'. But he became dissatisfied with this view because it "...involves the necessity of a definition of what is meant by a 'case'..." (Boole, 1854, ch. XI, §16), and Boole thinks that will involve us in all sorts of matters beyond the bounds of logic. So Boole takes a different view in *The Laws of Thought*:

> Let us take, as an instance for examination, the conditional proposition "If the proposition X is true, the proposition Y is true". An undoubted meaning of this proposition is, that the *time* in which the proposition X is true, is *time* in which the proposition Y is true. (Boole, 1854, ch. XI, §5)

The letters thus denote classes of times: the times a proposition is true. So the conditional proposition becomes a universal affirmative proposition: All times at which X is true are times at which Y is true. This is genius at its most twisted.

[46] We shall not pursue the matter here, but the fact that extensions of concepts are so critical to the Boolean tradition makes it a natural suggestion that Frege's own interest in extensions, which appear first in *Die Grundlagen* (Gl, §68), was also due to his encounter with Boole. As we noted earlier, the notion of unsaturatedness emerges at about the same time, and that is probably what enforced, for Frege, a distinction between concepts and extensions (cf. Heck, 2019).

[47] In fact, there is vigorous debate among the Booleans about whether '$A + B$' should be interpreted as union, which corresponds to inclusive disjunction, or to something rather more complicated that corresponds to exclusive disjunction. These disagreements are of no significance for the present discussion, however, so we shall ignore them.

[48] MacColl (1877, pp. 9–10) comes closest to this conception, but his official view is that the sentence-letters denote 'statements'. Schröder mentions MacColl in his review, but it is unclear if Frege ever read him. Frege mentions MacColl twice (AimCN, p. 93; BLF, p. 15), but what he says is all but lifted from Schröder.

Boole's reduction of the calculus of judgements to the calculus of classes thus amounts to his treating the sentential operators "$A \times B$", "$A + B$", and "$\bar{A}$" as expressing set-theoretic operations on the power set of some universe, indeed, as expressing the very same set-theoretic operations that "$A \times B$", "$A + B$", and "$\bar{A}$" express in the calculus of classes. The controversy about what comprises the universe has proved not to be the crucial point. On the contrary, the flexibility inherent in Boole's approach has proven a positive advantage. Boole's original view, which makes reference to 'cases', inspired some of the earliest work on modal logic, and his later view, which makes reference to times, had a similar influence on tense logic. Indeed, Boole's great insight is precisely that, *no matter what* we take the universe to comprise, if we treat the sentential operators as expressing set-theoretic operations on its power set, then the algebra so determined is (what we now call) a Boolean algebra, and it validates the laws of classical logic.

Blinded by the uncomprehending criticisms of his opponents, Frege could no more see the importance of this discovery of Boole's than Schröder could see the importance of Frege's notion of generality. But we must nonetheless agree with Frege that Boole's attempt to reduce sentential logic to quantification theory is a failure, and not only for the case of "eternal truths such as those of mathematics" (BLC, p. 15). What is fundamental is *sentential* logic, and Frege goes so far as to describe himself as reducing Boole's 'primary propositions', such as universal affirmative propositions, to his 'secondary' propositions, such as conditionals (BLC, pp. 17–8).

It is worth pausing over this point, because it reveals something important about Frege's early understanding of generality. What Frege means is that, in his logic, a statement like "All Fs are G" becomes a conditional (AimCN, p. 95):

$$\begin{array}{l} \vdash\!\!\!\top\, Gx \\ \;\;\llcorner\, Fx \end{array} \qquad (*)$$

Note carefully the omission of the leading quantifier. Frege could not have claimed to have reduced the calculus of classes to the calculus of judgements if he had symbolized "All Fs are G" as:

$$\begin{array}{l} \vdash\!\overset{\mathfrak{a}}{\smile}\!\!\top\, G\mathfrak{a} \\ \;\;\;\;\llcorner\, F\mathfrak{a} \end{array} \qquad (**)$$

Nor would it have served Frege's purposes to symbolize it as in (*) were this simply short for (**), as it would be if free variables were understood as tacitly bound by invisible universal quantifiers, as is often supposed. How exactly Frege understands free variables in *Grundgesetze* is a matter of some delicacy (see Heck, 2012a, §3.2), but it is clear how he understood them in *Begriffsschrift*. Frege's view at that time was that generality is indicated by "letters", that is, by variables (Bg, §1). This contrasts with his mature view, where generality is expressed by the 'concavity', that is, by the universal quantifier (which denotes a second-level concept that is true of precisely those concepts that are true of all). By contrast, Frege tells us in *Begriffsschrift* that the concavity serves only to "delimit[] the scope that the generality *indicated by the letter* covers" (Bg, §11, our emphasis). The concavity, that is to say, is merely an indicator of scope, a piece of pure syntax; it has no meaning of its own, and it is not a quantifier in the modern sense of that term.[49]

There is much more that could be said about this matter, but just two points are important at present. The first is that Boole does *not* treat the sentential connectives as expressing truth-functions. To be sure, Boole recognizes the importance of the special case in which the universe contains just one element. The calculations in which Boole is interested are especially simple in this case, which is of course that of a two-element Boolean algebra, with elements Boole would have denoted 1 and 0. But Boole simply does not interpret 1 and 0 as truth and falsity: They denote the universe and the empty set. And the same seems to have been true of Schröder, at least at the time he wrote his review of *Begriffsschrift*. What Schröder (1972, p. 224) says about sentential logic in the review is in agreement with what was quoted from Boole above.

So Boole did not treat the sentential connectives as expressing truth-functions, so Frege did not get the idea that they do from Boole. But Boole

[49] We have discussed this point in more detail elsewhere (Heck and May, 2013, §28.2), and it has been noted by others, too. We have a dim memory of having encountered it in the work of Peter Geach, but we cannot locate a reference. Gary Kemp (1995, p. 46, n. 12) mentions it in a footnote, but takes it to have been Frege's view throughout his career. That is clearly wrong: Frege explicitly says that in *Grundgesetze* that the quantifers have a reference (Gg, v. I, §31), and he tells us explicitly what that is (Gg, v. I, §§8, 24). What Kemp actually discusses in the footnote cited is Frege's understanding of Roman letters (i.e., free variables), which is a different matter. But, for what it is worth, we do not think Frege thought of free variables as "conferring generality" in his mature period either, though, as we have said, that matter is more delicate. The passage from "Introduction to Logic" (IntLog, p. 190) that Kemp cites cannot establish the point on its own. Frege's point there is that, in something like (*), the variable takes scope over the whole formula, not separately over its individual parts. This point could as well have been made with bound variables, though these have not yet been introduced at that point in the paper.

undoubtedly *did* treat the connectives as expressing *functions*. Boole's use of the arithmetical expressions "$+$" and "$\times$" serves to emphasize this point. And, as critical as Frege is of Boole's over-reliance on the analogy with arithmetic, it is hard to imagine that Frege would not have been impressed by this element of Boole's approach. It might even have seemed to Frege a confirmation of his own emphasis on the importance of the notion of function to logic. Frege does not highlight this aspect of Boole's work in the critical pieces from 1881 and 1882—he is too busy defending himself—but nor does he criticize it. So, although it may have taken him a little while to assimilate it, it seems very plausible that Frege *did* get the idea that the sentential connectives express functions from his reading of Boole.

The significance of this transformation should not be underestimated, quite independently of any correlative commitment to the *truth*-functionality of the connectives. It constitutes Frege's final abandonment of the traditional conception of logic as purely formal. That sentential connectives express functions makes them, for Frege, of a piece not only with the arithmetical expressions that belong to the 'material' part of the language but with predicates quite generally.[50] So, although the distinction between the logical and non-logical parts of the language will survive—Frege certainly would have regarded geometrical primtives as non-logical, for example—that distinction can no longer be understood as one between the formal and the contentful. And that, of course, is all to the good, so far as the requirements of Frege's logicism are concerned.[51]

Our present concern, however, is with the emergence of Frege's views about truth. So the next question we must ask, on Frege's behalf, is the one that must naturally arise once we have decided to treat the sentential connectives as expressing functions, namely:[52] What are the arguments and values of these functions? At one time, of course, the obvious thing for Frege to say would have been that the conditional expresses a two-place function from contents to contents. But this option is off the table now, due to the antinomy we discussed in Section 4.

[50] It is this aspect of Frege's view that Wittgenstein targets in the *Tractatus* (Wittgenstein, 1961). For discussion, see May (2018, esp.\{} pp.\{} 113–6).

[51] Vestiges of this traditional view seem to survive at least into Quine (1970), whose opposition to the logical status of second-order logic is otherwise hard to understand. Or, at least some of his grounds are hard to understand. Why should it matter otherwise if second-order logic admits of a complete proof procedure? Quine seems to be supposing, quite without argument, that logic is essentially syntactic (cf. Boolos, 1975, pp. 525–6).

[52] As said earlier, this question will only arise once Frege has stopped confusing functions with expressions. But we are past that point now.

Note first that the arguments of the sentential functions must be of the same sort as their values: Since negation and the conditional embed within one another, the value of the one becomes an argument of the other. And the arguments of the connectives will plausibly be the same as the values of concept-functions, too, since, in a formula like

$$\begin{array}{l} \vdash\!\!\top\, Fb \\ \;\;\llcorner\, Fa \end{array}$$

the values of the concept-function $F\xi$ become the arguments of the conditional.[53] We know what Frege's mature view was: All of these are truth-values. But what leads him to that view?

Recall Frege's explanation of the conditional in *Begriffsschrift*:

> If A and B stand for contents that can become judgements. . . , there are the following four possibilities:
>
> (1) A is affirmed and B is affirmed;
>
> (2) A is affirmed and B is denied;
>
> (3) A is denied and B is affirmed;
>
> (4) A is denied and B is denied.
>
> Now
>
> $$\begin{array}{l} \vdash\!\!\top\, A \\ \;\;\llcorner\, B \end{array}$$
>
> stands for the judgement that the third of these possibilities does not take place, but one of the other three does. (Bg, §5, emphasis removed)

The language of affirmation and denial is not only quaint but misplaced, as Frege himself would eventually come to realize: When one asserts a conditional, one is *not* thereby affirming or denying its antecedent or consequent, and one is not saying anything about whether anyone else should affirm or deny them, either. This is essentially what Peter Geach (1965) famously called "the Frege point", and it is closely connected with

[53] None of this is absolutely necessary, due to the presence of the horizontal, or content-stroke, in Frege's language. If concept-functions had conceptual contents for their values (or circumstances, or whatever), then "$-\!\!-\, \xi$" might yet be understood as expressing a function from conceptual contents to truth-values—as in effect being a truth-predicate—and then everything would still fit together. It is even possible that Frege at some time held this sort of position, though we know of nowhere that he expresses it.

what Frege himself called "the dissociation of assertoric force from the predicate" (IntLog, p. 185).

The crucial observation is, yet again, that Frege changes his view about this matter. There are certainly intimations of the Frege point in *Begriffsschrift*, e.g, in §2. But it is not, contrary to what Geach (1965, p. 449) claims, ever actually stated there. If Frege had fully appreciated the Frege point at that time, he could not have explained the conditional as he does, in terms of affirmation and denial. Such language continues to appear in the writings on Boole from 1881 and 1882, especially "Boole's Logical Formula-language" (BLF, esp. p. 50), but it does not appear later. So, at some point between 1882 and 1884—when, as we have seen, the Truth-Value Thesis is already in place—Frege must have become dissatisfied with his use of such language, probably for the reason we have just recalled.

When he does abandon the language of affirmation and denial, however, Frege is faced with a problem: He needs a new way to explain the conditional; he needs, in particular, to reconstruct his table of possibilities. In the writings from 1881 and 1882, Frege sometimes presents the table this way (BLC, p. 35; BLF, p. 49):

(1) A and B;

(2) A and not-B;

(3) not-A and B;

(4) not-A and not-B.

So long as one is not thinking of the conditional as expressing a function, then this will do. But once Frege has decided that the conditional ought to be understood as expressing a function, this form of explanation will no longer do, for now we must ask what the arguments and values of the function expressed by the conditional are, and simply listing the cases in this way does not answer that question. If that is the question, however, then it is not hard to see how, after his encounter with Boole, Frege might eventually have been struck by the answer: What is relevant is not whether A and B are affirmed or denied, but just whether they are true or false, and the arguments and values of the sentential connectives are 'truth-values'.[54]

[54] There is a complication here, for which see note 42.

Once this idea is in place, Frege's other problem gets resolved, too. As said previously, the arguments of the connectives are the values of the concept-functions, since, in a formula like

$$\begin{array}{l} \vdash\!\!\!\top\, Fb \\ \quad\llcorner\, Fa \end{array}$$

the values of the concept-function $F\xi$ become the arguments of the conditional. So concepts are functions from objects to truth-vaules. The fact that the elements of this package fit so nicely together then becomes, for Frege, yet another reason to think it must be right.

In the end, then, Frege's argument for the view that sentences refer to truth-values is, as we said earlier, broadly pragmatic: It solves a lot of problems. Taking sentences to refer to truth-values tells us what the values of concept-functions are and it gives us the powerful notion of truth-functionality. In short, the idea that sentences refer to truth-values ties up a bunch of loose ends. So it is no surprise that Frege should boast in *Grundgesetze* of "how much simpler and more precise everything is made by the introduction of truth-values" (Gg, v. I, p. x). Though he might have added: It would help to look at my earlier work so as to get a sense for how messy things used to be.

6 Closing

Our story of how truth came to play a central role in Frege's philosophy of logic is now complete. But, as we noted at the end of Section 4, Frege's position still is not quite stable. Taking sentences to refer to truth-values tells us how we can weaken "$\equiv$" so that

$$\vdash c = d \to Fc \equiv Fd$$

might be true but harmless. Indeed, it can now be reformulated simply as

$$\vdash c = d \to Fc = Fd$$

as it is in Frege's mature work, with an identity between sentences being true just in case they have the same truth-value. But other problems loom, as Frege himself notes in "On Sense and Reference":[55]

[55] We here translate "*Bedeutung*" as meaning, rather than reference. One simply cannot appreciate the force of the worry otherwise: that all true sentences *mean* the same thing.

> If now the truth-value of a sentence is its meaning, then on the one hand all true sentences have the same meaning and so, on the other hand, do all false sentences. From this we see that in the meaning of the sentence all that is specific is obliterated. (SM, op. 35)

And the problem is not just that this might seem strange. The problem is that we now have no way to understand the nature of inference.

As we have emphasized, an inference, for Frege, is a transition between judgements, and the correctness of such an inference is supposed to be determined by the contents of those judgements: Sameness of content is supposed to imply sameness of logical relations; differences of content are to be recognized only where the logical relations are different. But if "in the meaning of the sentence all that is specific is obliterated", then content can be meaning only if all true sentences, and all false sentences, have the same logical properties, which is absurd. There must, therefore, be another sort of content, one that is more fine-grained, in terms of which the correctness of inferences can be characterized.

The problem with which the Truth-Value Thesis leaves Frege is thus one that is utterly central to his conception of logic. Frege wanted to insist that logic is, in a sense, formal, but not in the sense that logic is concered only with form. On the contrary, the task of logic is to uncover the laws of truth, and there is no truth without content. More precisely, inferences of the form

$$A \to B, A \vdash B$$

are, Frege wants to insist, valid not in virtue of formal relations between the symbols that appear in these formulae but in virtue of semantic relations between the contents expressed by the formulae. Logic is formal only in so far as formal relations between symbols can be made to reflect semantic relations between contents.

The problem that animates Frege's mature philosophy is how this idea might be reconciled with seemingly incompatible demands. On the one hand, Frege's insistence that logic must concern itself with truth leads him to say that the conditional itself expresses a truth-function, one that will not permit B to be false if both A and $A \to B$ are true. That is why the inference is valid. But, on the other hand, if the conditional expresses a function, then, once it is given some arguments, it will, so to speak, apply itself to them and hand us a value. From the point of view of meaning, then, all we have is

$$\top, \top \vdash \top$$

and nothing about *modus ponens* follows from that triviality. What we seem to need is a way of holding the arguments apart from the function, so that they are not obliterated by functional application, obscuring the nature of the inference. We need, that is to say, is to find some way to retain the structural features of a sentence in its content, and the broader lesson of the substitution argument is that the notions of function and argument are ill-suited to this purpose.[56]

Frege's proposal, in his mature work, is that we should regard A and B as expressing 'thoughts' that are themselves components of the thought expressed by "$A \to B$". But this only leaves us with several more questions. The most obvious is how exactly the thought expressed by the conditional is supposed to be composed from the thoughts expressed by its parts. We have discussed that vexed question elsewhere (Heck and May, 2010). But there is an even more pressing question. It was in terms of a relation between their truth-values that we explained the validity of the inference from the judgement that $A \to B$ and the judgement that A to the judgement that B. If we are still to endorse such an explanation, then we need to understand how this relation among the truth-values of these thoughts depends upon the structural relations among them. More generally, we need to understand how thoughts are related to their truth-values.[57] No relation between thoughts can make an inference valid unless there is some such relation—not if inference is to have anything to do with truth or, better, with knowledge, as both Frege's logicism and common sense require.

Frege's solution to this problem is that we should regard the relation of a thought to its truth-value as that of sense to reference. What this might mean, unfortunately, is still far from clear. Part of what it means, or so we would venture to suggest, is that, as Dummett (1993c) would have put it, truth is the central notion of semantics. But how exactly

[56] This is so even if we move the function–argument analysis up to the level of sense, taking the sense of a predicate to be a function from senses to thoughts (Heck and May, 2010, p. 147).

[57] Eva Picardi (1996, p. 318) puts a closely related point thus:

> Even a superficial reader of [*Grundgesetze*] soon realizes that without the account of how a sentence of the formal language is determined as true—the sense expressed by an assertion of a sentence of [*Grundgesetze*] being the thought that its truth-conditions are fulfilled [(Gg, v. I, §32)]—Frege would have not a shade of an argument for opposing the doctrines both of psychologistic logicians and of formalist mathematicians.

An "account of how a sentence of the formal language is determined as true" is of course a semantics.

Frege's idea should be developed is a question we shall have to leave for another time.[58]

References

Austin, J. L. (1951). 'Truth', *Proceedings of the Aristotelian Society* 51: 111–28.

Bar-On, Dorit and Simmons, Keith (2007). 'The use of force against deflationism: Assertion and truth', in Dirk Greimann and Geo Siegwart (eds.), *Truth and Speech Acts: Studies in the Philosophy of Language*. New York, Routledge, 61–89.

Beaney, Michael (2007). 'Frege's use of function-argument analysis and his introduction of truth-values as objects', *Grazer Philosophische Studien* 75: 93–123.

Boole, George (1847). *The Mathematical Analysis of Logic*. Cambridge, Macmillan, Barclay, & Macmillan.

—— (1854). *The Laws of Thought*. London, Walton and Maberly.

Boolos, George (1975). 'On second-order logic', *Journal of Philosophy* 72: 509–27.

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

Bynum, Terrell Ward (1972). 'Editor's introduction', in Frege 1972, 55–80.

Davidson, Donald (1984). 'Truth and meaning', in *Inquiries Into Truth and Interpretation*. Oxford, Clarendon Press, 17–36.

—— (1990). 'The structure and content of truth', *Journal of Philosophy* 87: 279–328.

[58] This paper began life as a draft of what would eventually become our survey "Truth In Frege" (Heck and May, 2018). Unfortunately, the paper quickly exceeded the word limit we had been given, so we had to excerpt and re-organize what seemed most appropriate for the *Oxford Handbook of Truth*. Thanks to Michael Glanzberg for the invitation to write that piece (and so for the inspiration for this one) and to an anonymous referee for helpful comments.

—— (2005). ‘The folly of trying to define truth’. Oxford, Clarendon Press, 19–37.

Dummett, Michael (1978). ‘Truth’, in *Truth and Other Enigmas*. London, Duckworth, 1–24.

—— (1981). *Frege: Philosophy of Language*, 2d edition. Cambridge MA, Harvard University Press.

—— (1991). *The Logical Basis of Metaphysics*. Cambridge MA, Harvard University Press.

—— (1993a). *The Seas of Language*. Oxford, Clarendon Press.

—— (1993b). ‘The source of the concept of truth’, in Dummett 1993a, 188–201.

—— (1993c). ‘What is a theory of meaning? (II)’, in Dummett 1993a, 34–93.

Fine, Kit (2007). *Semantic Relationism*. Oxford, Blackwell.

Frege, Gottlob (1964). *The Basic Laws of Arithmetic: Exposition of the System*, ed. and tr. by Montgomery Furth. Berkeley CA, University of California Press.

—— (1970). *Translations from the Philosophical Writings of Gottlob Frege*, P. T. Geach and M. Black, eds. Oxford, Blackwell.

—— (1972). *Conceptual Notation and Related Articles*, T. W. Bynum, ed. New York, Oxford University Press.

—— (1979). *Posthumous Writings*, H. Hermes, F. Kambartel, and F. Kaulbach, eds. Chicago, University of Chicago Press.

—— (1984). *Collected Papers on Mathematics, Logic, and Philosophy*, B. McGuiness, ed. Oxford, Basil Blackwell.

—— (1997). *The Frege Reader*, M. Beaney, ed. Oxford, Blackwell.

—— (AimCN). ‘On the aim of the conceptual notation’, tr. by T. W. Bynum, in Frege 1972, 90–100.

—— (Bg). 'Begriffsschrift: A formula language modeled upon that of arithmetic, for pure thought', tr. by S. Bauer-Mengelberg, in Jean van Heijenoort (ed.), *From Frege to Gödel: A Sourcebook in Mathematical Logic 1879–1931*. Cambridge MA, Harvard University Press, 5–82.

—— (BLC). 'Boole's logical calculus and the concept-script', in Frege 1979, 9–46.

—— (BLF). 'Boole's logical formula-language and my concept-script', in Frege 1979, 47–52.

—— (CO). 'On concept and object', tr. by P. T. Geach, in Frege 1984, 182–94. Also in Frege 1997, 181–93.

—— (CSM). 'Comments on sense and meaning', in Frege 1979, 118–25. Also in Frege 1997, 172–80.

—— (CT). 'Compound thoughts', tr. by P. T. Geach and R. H. Stoothoff, in Frege 1984, 390–406.

—— (FC). 'Function and concept', tr. by P. T. Geach, in Frege 1984, 137–56. Also in Frege 1997, 130–48.

—— (FTA). 'Formal theories of arithmetic', tr. by E.-H. W. Kluge, in Frege 1984, 112–21.

—— (Geo2). 'On the foundations of geometry: Second series', tr. by E.-H. W. Kluge, in Frege 1984, 293–340.

—— (Gg). *The Basic Laws of Arithmetic*, tr. by P. A. Ebert and M. Rossberg. Oxford, Oxford University Press.

—— (Gg). *Grundgesetze der Arithmetik*. Hildesheim, Georg Olms Verlagsbuchhandlung. Translations of Part I are in Frege 1964, and of portions of Part III, in Frege 1970. Extracts appear in Frege 1997, 194–223.

—— (Gl). *The Foundations of Arithmetic*, 2d revised edition, tr. by J. L. Austin. Evanston IL, Northwestern University Press.

—— (IntLog). 'Introduction to logic', in Frege 1979, 185–96. Extracts in Frege 1997, 293–8.

—— (Log79). 'Logic (1879–1891)', in Frege 1979, 1–8.

—— (Log97). ‘Logic (1897)’, in Frege 1979, 126–151. Extracts in Frege 1997, 227–50.

—— (MBLI). ‘My basic logical insights’, in Frege 1979, 251–252. Also in Frege 1997, pp. 322–4.

—— (PCN). ‘On Mr. Peano’s conceptual notation and my own’, tr. by V. H. Dudman, in Frege 1984, 234–48.

—— (PMC). *Philosophical and Mathematical Correspondence*, G. Gabriel, H. Hermes, F. Kambartel, C. Thiel, and A. Veraart, eds., tr. by H. Kaal. Chicago, University of Chicago Press.

—— (Schr). ‘A critical elucidation of some points in E. Schröder, *Vorlesungen über die Algebra der Logik*’, in Frege 1984, 210–28.

—— (SM). ‘On sense and meaning’, tr. by M. Black, in Frege 1984, 157–77. Also in Frege 1997, 151–71.

—— (Tht). ‘Thoughts’, tr. by P. T. Geach and R. H. Stoothoff, in Frege 1984, 351–72. Also in Frege 1997, 325–45.

Geach, P. T. (1965). ‘Assertion’, *Philosophical Review* 74: 449–65.

Greimann, Dirk (2015). ‘Frege’s performative argument against truth relativism’, *Journal for the History of Analytic Philosophy* 3(2).

Heck, Richard Kimberly (2012a). *Reading Frege’s Grundgesetze*. Oxford, Clarendon Press. Originally published under the name ‘Richard G. Heck, Jr’.

—— (2012b). ‘Solving Frege’s puzzle’, *Journal of Philosophy* 109: 132–74. Originally published under the name ‘Richard G. Heck, Jr’.

—— (2014). ‘Intuition and the substitution argument’, *Analytic Philosophy* 55: 1–30. Originally published under the name ‘Richard G. Heck, Jr’.

—— (2019). ‘Formal arithmetic before *Grundgesetze*’, in Philip Ebert and Marcus Rossberg (eds.), *Essays on Frege’s Basic Laws of Arithmetic*. Oxford, Oxford University Press, 497–537. Forthcoming.

Heck, Richard Kimberly and May, Robert (2006). ‘Frege’s contribution to philosophy of language’, in Ernest Lepore and Barry Smith (eds.), *The Oxford Handbook of Philosophy of Language*. Oxford, Oxford University Press, 3–39. Originally published under the name 'Richard G. Heck, Jr'.

—— (2010). 'The composition of thoughts', *Noûs* 45: 126–66. Originally published under the name 'Richard G. Heck, Jr'.

—— (2013). 'The function is unsaturated', in Michael Beaney (ed.), *The Oxford Handbook of the History of Analytic Philosophy*. Oxford, Oxford University Press, 825–50. Originally published under the name 'Richard G. Heck, Jr'.

—— (2018). 'Truth in Frege', in Michael Glanzberg (ed.), *The Oxford Handbook of Truth*. Oxford, Oxford University Press, 193–215.

Kemp, Gary (1995). 'Truth in Frege's "law of truth"', *Synthese* 105: 31–51.

Kneale, William and Kneale, Martha (1962). *The Development of Logic*. Oxford, Clarendon Press.

Kremer, Michael (2010). 'Sense and meaning: The origins and development of the distinction', in Michael Potter and Thomas Ricketts (eds.), *The Cambridge Companion to Frege*. Cambridge, Cambridge University Press, 220–92.

Larson, Richard and Segal, Gabriel (1995). *Knowledge of Meaning*. Cambridge MA, MIT Press.

Linnebo, Øystein (2003). 'Frege's conception of logic: From Kant to *Grundgesetze*', *Manuscrito* 26: 235–52.

MacColl, Hugh (1877). 'The calculus of equivalent statements and integration limits', *Proceedings of the London Mathematical Society* IX: 9–20.

—— (1880). 'Symbolical reasoning', *Mind* 5: 45–60.

Marcus, Ruth Barcan (1995). 'Modalities and intensional languages', in *Modalities: Philosophical Essays*. Oxford, Oxford University Press, 3–24.

May, Robert (2001). 'Frege on identity statements', in C. Cecchetto, *et al*. (eds.), *Semantic Interfaces: Reference, Anaphora, and Aspect*. Stanford, CSLI Publications, 1–62.

—— (2006). ‘The invariance of sense’, *Journal of Philosophy* 103: 111–144.

—— (2012). ‘What frege’s theory of identity is not’, *Thought* 1: 41–48.

—— (2018). ‘Logic as science’, in Annalisa Coliva, *et al*. (eds.), *Eva Picardi on Language, Analysis and History*. London, Palgrave Macmillan, 113–60.

Moore, G. E. (1953). *Some Main Problems of Philosophy*. London, Allen and Unwin.

Pedriali, Walter B. (2017). ‘The logical significance of assertion: Frege on the essence of logic’, *Journal for the History of Analytic Philosophy* 5 (8).

Picardi, Eva (1996). ‘Frege’s anti-psychologism’, in Matthias Schirn (ed.), *Frege: Importance and Legacy*. New York, de Gruyter, 307–29.

Post, Emil (1921). ‘Introduction to a general theory of elementary propositions’, *The American Journal of Mathematics* 43: 163–85.

Proops, Ian (1997). ‘The early Wittgenstein on logical assertion’, *Philosophical Topics* 25: 121–44.

Quine, W. V. O. (1953). ‘On what there is’, in *From a Logical Point of View*. Cambridge MA, Harvard University Press, pp. 1–18.

—— (1970). *Philosophy of Logic*. Englewood Cliffs NJ, Prentice Hall.

Ricketts, Thomas (1996). ‘Logic and truth in Frege’, *Proceedings of the Aristotelian Society* sup. vol. 70: 121–40.

—— (1997). ‘Frege’s 1906 foray into meta-logic’, *Philosophical Topics* 25: 169–88.

Russell, Bertrand (1903). *The Principles of Mathematics*. Cambridge, Cambridge University Press.

Schröder, Ernst (1877). *Der Operationskreis des Logikkalküls*. Leipzig, B.G. Teubner.

—— (1972). ‘Review of Frege’s *Conceptual Notation*’, in Frege 1972, 218–32.

Strawson, P. F. (1950). ‘Truth’, *Proceedings of the Aristotelian Society* supl vol 24: 129–56.

Tarski, Alfred (1956). ‘The concept of truth in formalized languages’, in John Corcoran (ed.), *Logic, Semantics, and Metamathematics*. Indianapolis, Hackett, 152–278.

Textor, Mark (2010). ‘Frege on judging as acknowledging the truth’, *Mind* 119: 615–55.

Venn, John (1894). *Symbolic Logic*, 2d edition. London, MacMillan.

—— (1972). ‘Review of Frege’s *Conceptual Notation*’, in Frege 1972, 234–5.

Wiggins, David (1980). ‘What would be a substantial theory of truth?’, in Zaak van Straaten (ed.), *Philosophical Subjects*. Oxford, Clarendon Press, 189–221.

Wittgenstein, Ludwig (1961). *Tractatus Logico-Philosophicus*. London, Routledge and Kegan Paul.

—— (1979a). *Notebooks: 1914–1916*, 2d edition, G. H. von Wright and G. E. M. Anscombe, eds. Chicago, University of Chicago Press.

—— (1979b). *Wittgenstein’s Lectures: Cambridge 1932–1935*, A. Ambrose, ed. Chicago, University of Chicago Press.