

## What's Puzzling Gottlob Frege?<sup>1</sup>

MIKE THAU and BEN CAPLAN  
University of California, Los Angeles  
Los Angeles, CA 90095  
USA

### I Introduction

By any reasonable reckoning, Gottlob Frege's 'On Sense and Reference'<sup>2</sup> is one of the more important philosophical papers of all time. Although Frege briefly discusses the sense-reference distinction in an earlier work ('Function and Concept,' in 1891), it is through 'Sense and Reference' that most philosophers have become familiar with it. And the distinction so thoroughly permeates contemporary philosophy of language and mind that it is almost impossible to imagine these subjects without it.

1 For support in the form of a Doctoral Fellowship (#752-98-0363), thanks from the second author to the Social Sciences and Humanities Research Council (SSHRC) of Canada. For comments and discussion, thanks from both authors to the editors of, and anonymous referees for, the *Canadian Journal of Philosophy* and also to Kent Bach, Glenn Branch, Tyler Burge, Louis Derossset, David Kaplan, Sean Kelsey, Robert May, Scott Soames, and especially Mark Kalderon.

2 'Über Sinn und Bedeutung,' *Zeitschrift für Philosophie und philosophische Kritik* 100 (1892) 25-50. The essay was translated by Max Black as 'On Sense and Meaning' in *Translations from the Philosophical Writings of Gottlob Frege*, 3rd ed., Peter T. Geach and Max Black, eds. (Oxford: Blackwell 1980), 56-78. It was reprinted as 'On Sinn and Bedeutung' in *The Frege Reader*, Michael Beaney, ed. (Oxford: Blackwell 1997), 151-71.

Black translates 'Bedeutung' as 'meaning' and 'bedeuten' as 'mean' (or sometimes as 'designate'); Beaney translates 'bedeuten' as 'stand for' (or sometimes as 'denote'). We emend their translations, of 'Sense and Reference' and other texts, so that 'Bedeutung' is translated as 'reference' or 'referent' and 'bedeuten' is translated as 'refer to,' because we find this translation the most familiar and the easiest to understand. Other emendations to previous translations are indicated in the notes.

The distinction between the sense and the referent of a name is introduced in the second paragraph of 'Sense and Reference'.<sup>3</sup> The first paragraph lays the groundwork for its introduction and runs as follows:

Equality gives rise to challenging questions which are not altogether easy to answer. [Footnote to 'Equality': I use this word in the sense of identity and understand 'a = b' to have the sense of 'a is the same as b' or 'a and b coincide.'] Is it a relation? A relation between objects, or between names or signs of objects? In my *Begriffsschrift* I assumed the latter. The reasons which seem to favor this are the following: a = a and a = b are obviously statements of differing cognitive value; a = a holds *a priori* and, according to Kant, is to be labeled analytic, while statements of the form a = b often contain very valuable extensions of our knowledge and cannot always be established *a priori*. The discovery that the rising sun is not new every morning, but always the same, was one of the most fertile astronomical discoveries. Even today the reidentification of a small planet or comet is not always a matter of course. Now if we were to regard equality as a relation between that which the names "a" and "b" refer to, it would seem that a = b could not differ from a = a, i.e. provided that a = b is true. A relation would thereby be expressed of a thing to itself, and indeed one in which each thing stands to itself but to no other thing. What we apparently want to state by a = b is that the signs or names "a" and "b" refer to the same thing, so that those signs themselves would be under discussion; a relation between them would be asserted. But this relation would hold between the names or signs only in so far as they named or designated something. It would be mediated by the connection of each of the two signs with the same designated thing. But this is arbitrary. Nobody can be forbidden to use any arbitrarily producible event or object as a sign for something. In that case the sentence a = b would no longer touch the subject matter,<sup>4</sup> but only its mode of designation; we would express no proper knowledge by its means. But in many cases this is just what we want to do. If the sign "a" is distinguished from the sign "b" only as an object (here, by means of its shape), not as a sign (i.e. not by the manner in which it designates something), the cognitive value of a = a becomes essentially equal to that of a = b, provided that a = b is true. A difference can arise only if the difference between the signs corresponds to a difference in the mode of presentation of the thing designated. Let a, b, c be the lines connecting the vertices of a triangle with the midpoints of the opposite sides. The point of intersection of a and b is then the same as the point of intersection of b and c. So we have different designations for the same point, and these names ("point of intersection of a and b," "point of intersection of b and c") likewise indicate the mode of presentation; and hence the statement contains actual knowledge. ('On Sense and Reference,' 25-6; 25 n.A)

3 We use 'name' (or 'sign') as Frege does to include all singular terms. See 'Sense and Reference,' 27.

4 We translate 'betreffen' here as 'touch' (rather than 'be concerned with'), as it is translated in section 8 of the *Begriffsschrift*; and we translate 'Sache selbst' there as 'subject matter' (rather than 'matter'), as it is translated here. On our view, there is a close connection between these passages, so it is important to bring their translations in line.

In the next paragraph, Frege introduces the notion of *sense* by saying that the sense of a name is 'wherein the mode of presentation is contained' (26-7).

As with any greatly significant philosophical work, there is much disagreement about the details of Frege's discussion. Nonetheless, there is something of a standard interpretation of what goes on in the opening paragraph. Indeed, it is an interpretation that both of us were taught as undergraduates and have, until recently, taught our own students. According to the standard interpretation, Frege's discussion in the opening paragraph runs as follows:<sup>5</sup>

- (i) Frege begins by wondering which of two views of identity is correct. On the first, metalinguistic view — which we'll call the *name view* — [ $\alpha = \beta$ ] expresses the thought that the names  $\alpha$  and  $\beta$  have the same referent; whereas, on the second, non-metalinguistic view — which we'll call the *object view* — [ $\alpha = \beta$ ] expresses the thought that a and b are the same object, where a and b are the objects named by  $\alpha$  and  $\beta$ , respectively. Frege states that in an earlier work, his *Begriffsschrift*,<sup>6</sup> he assumed that the name view was correct.
- (ii) Frege gives his reason for choosing the name view over the object view in the *Begriffsschrift*. His reason is *not* that the name view has some virtue, but rather that the object view has some vice: namely, that on the object view [ $\alpha = \alpha$ ] and [ $\alpha = \beta$ ] (if true) turn out to express the same thing — each expresses, of the very same object, that it is identical with itself.
- (iii) Frege goes on to point out that the name view also has a vice: namely, that on the name view all identity statements turn out to

Dummett

5 Dummett, Salmon, Sluga, and Weiner endorse the standard interpretation. See Michael Dummett, *Frege: Philosophy of Language*, 2nd ed. (Cambridge, MA: Harvard University Press 1981), 544; Nathan Salmon, *Frege's Puzzle* (Cambridge, MA: The MIT Press 1986), 51-4; Hans Sluga, *Gottlob Frege* (Boston: Routledge & Kegan Paul 1980), 149-61; and Joan Weiner, *Frege* (Oxford: Oxford University Press 1999), 91-2. Kaplan agrees that Frege rejects the name view, but he does not say that Frege accepts the object view instead. See David Kaplan, 'Words,' *Supplement to the Aristotelian Society* 64 (1990) 93-119, at 118.

6 *Begriffsschrift, eine der arithmetischen nachgebildete Formelsprache des reinen Denkens* (Halle: Nebert 1879). Preface and Part I trans. by Michael Beaney as 'Begriffsschrift: A Formula Language of Pure Thought Modelled on that of Arithmetic' in *The Frege Reader*, 47-78.

be about language. For example, on the name view the sentence 'Hesperus is Phosphorus' turns out to express that the names 'Hesperus' and 'Phosphorus' refer to the same object. But the sentence is about astronomy, not language, so the name view gets the subject matter of identity statements wrong.

- (iv) Frege rejects the name view, on which identity statements express metalinguistic thoughts, and reconsiders his earlier objection to the object view. He realizes that this objection can be met if names are associated with *modes of presentation* as well as *referents*.<sup>7</sup> If  $\lceil\alpha = \beta\rceil$  is true, then  $\alpha$  and  $\beta$  have the same referent, so  $\lceil\alpha = \alpha\rceil$  and  $\lceil\alpha = \beta\rceil$  both attribute identity to the same pair of objects. But the truth of  $\lceil\alpha = \beta\rceil$  does not imply that  $\alpha$  and  $\beta$  are associated with the same mode of presentation. And, if the modes of presentation associated with  $\alpha$  and  $\beta$  determine which thoughts  $\lceil\alpha = \alpha\rceil$  and  $\lceil\alpha = \beta\rceil$  express, then the two sentences can express different thoughts, even though  $\alpha$  and  $\beta$  have the same referent. For example, the mode of presentation associated with 'Hesperus' might be given by 'the evening star,' while that associated with 'Phosphorus' might be given by 'the morning star.' In that case, 'Hesperus is Hesperus' expresses the non-metalinguistic thought that the evening star is the evening star, while 'Hesperus is Phosphorus' expresses the non-metalinguistic thought that the evening star is

the morning star.<sup>8</sup> To put the point in contemporary terminology, the object view implies that  $\lceil\alpha = \alpha\rceil$  and  $\lceil\alpha = \beta\rceil$  express the same proposition *only* on the assumption that names are *directly referential*,<sup>9</sup> but if names instead refer to their referents via modes of presentation, then  $\lceil\alpha = \alpha\rceil$  and  $\lceil\alpha = \beta\rceil$  need not express the same proposition.

As the reader may have guessed by now, we have come to bury the standard interpretation rather than to praise it. We have no quarrel with (i)-(iii), but we do reject (iv). In particular, the main point of this paper is to argue (Frege scholars prepare to be surprised) that, contrary to (iv), Frege never rejects the view that identity statements express metalinguistic thoughts.<sup>10</sup> We suspect that many readers will find this unbelievable. After all, doesn't Frege say in the opening paragraph that he is rejecting his view from the *Begriffsschrift* that  $\lceil\alpha = \beta\rceil$  expresses that  $\alpha$  and  $\beta$  co-refer? Actually, although Frege does raise an objection to that metalinguistic view, nowhere in 'Sense and Reference' does he say that it is wrong. And, more importantly, a careful reading of the text reveals that Frege is up to something quite different than the standard interpretation supposes.

So we think that the text of 'Sense and Reference' *itself* gives the lie to the standard interpretation. But detailed textual analysis is, by its very

7 We do not draw any distinction between senses and modes of presentation. However, as stated above in the text, Frege claims in the second paragraph of 'Sense and Reference' that a sense *contains* a mode of presentation, which suggests that they are distinct. But the containment metaphor does not explain what, if anything, the distinction between senses and modes of presentation really amounts to. Moreover, after introducing the notion of a mode of presentation in 'Sense and Reference,' Frege gives examples of them; and, in each of his examples, a mode of presentation is given by a definite description: for example, by 'the intersection of lines l and m.' But, in a footnote to 'Sense and Reference,' Frege says: 'In the case of an actual proper name such as "Aristotle" opinions as to the sense may differ. It might, for instance, be taken to be the following: the pupil of Plato and teacher of Alexander the great' (27 n.B). So, in the footnote, Frege gives the *sense* (rather than the mode of presentation contained in the sense) of the name by a definite description. This makes it very doubtful that the distinction between modes of presentation and senses that Frege seems to implicitly draw in the text represents something of importance to him. Indeed, we think the footnote (coupled with the fact that Frege says absolutely nothing about the distinction beyond the metaphor) makes it plausible that the apparent implication of the containment metaphor may be unintentional — that Frege doesn't mean to be suggesting that modes of presentation and senses are distinct. In any case, as far as we can tell, whatever distinction there might be between them has no bearing on our discussion.

8 We do not mean to take a stand here on whether Frege is a descriptivist: that is, whether he thinks that the sense of *every* name is given by a definite description. But it is clear (see note 7 above) that Frege thinks that the senses of at least *some* names are given by definite descriptions.

9 The notion of a directly referential expression is developed in David Kaplan, 'Demonstratives: An Essay on the Semantics, Logic, Metaphysics, and Epistemology of Demonstratives and Other Indexicals,' *Themes from Kaplan*, Joseph Almog, John Perry, and Howard Wettstein, eds. (New York: Oxford University Press 1989), 481-563.

10 In the course of our research, we have come across various writers whose views, to various degrees, anticipate our own. See Avrum Stroll, 'Identity,' *Encyclopedia of Philosophy* vol. 4, Paul Edwards, ed. (New York: Macmillan 1967), 123-4; Jan Dejnozka, 'Frege on Identity,' *International Studies in Philosophy* 13.1 (1981) 31-41, at 31-6, and *The Ontology of the Analytic Tradition and its Origins: Realism and Identity in Frege, Russell, Wittgenstein, and Quine* (Lanham, MD: Littlefield Adams 1996), 42-65; Gregory Currie, *Frege: An Introduction to his Philosophy* (Brighton: Harvester 1982), 98-100 and 108-12; and Thomas V. Morris, *Understanding Identity Statements* (Aberdeen: Aberdeen University Press 1984), 27-34. But each falls short in some way. Discussing the details of Frege's texts is complicated enough; so, rather than explicitly signaling where we disagree with or go beyond other interpretations, we simply argue for our own.

nature, subtle; and, as we all know, subtleties seldom suffice to wipe out deeply ingrained views. As one philosopher pointed out in another context, our habitual opinions have a tendency to keep coming back and capturing our belief, which is, as it were, bound over to them as a result of long occupation and the law of custom. So, before turning to the evidence from 'Sense and Reference,' we consider evidence from other texts — evidence that has the advantage of being a good deal less subtle.

## II Frege's Continued Assertion of the Name View

The *Begriffsschrift* is not the only place where Frege asserts that identity statements express metalinguistic thoughts. In a few places before 'Sense and Reference,' Frege assumes the name view in arguing against a crude version of formalism, according to which mathematical expressions that seem to refer to numbers actually have no referent. First, in 'On Formal Theories of Arithmetic,' Frege says:

According to the above [formalist view], one also could not say that  $1/2$  is equal to  $3/6$ , for both are merely figures. The situation changes radically when one takes these figures to be signs of contents; in that case the equation states that both names have the same content.<sup>11</sup>

Frege is claiming that ' $1/2 = 3/6$ ' states that ' $1/2$ ' and ' $3/6$ ' refer to the same object and, hence, is clearly assuming that an identity statement states that the names that flank the identity sign have the same referent.<sup>12</sup>

11 'Über formale Theorien der Arithmetik,' *Sitzungsberichte der Jenaischen Gesellschaft für Medizin und Naturwissenschaft* 19 (Supplement 2, 1885) 94-104, at 101. Trans. by Eike-Henner W. Kluge as 'On Formal Theories of Arithmetic' in *On the Foundations of Geometry and Formal Theories of Arithmetic* (New Haven: Yale University Press 1971), 141-53.

12 Frege says 'have the same content' ('denselben Inhalt haben') rather than 'refer to the same object.' In 1885, Frege had yet to clearly distinguish sense and reference. He later came to think that sentences themselves are names and that his use of 'the content (*Inhalt*) of a sentence' in the *Begriffsschrift* was ambiguous between the sense of a sentence and its referent. See his letter to Husserl, 'Frege an Husserl 24.5.1891,' *Wissenschaftlicher Briefwechsel*, Gottfried Gabriel et al., eds. (Hamburg: Meiner 1976), 94-8, at 97; trans. by Hans Kaal as 'Frege to Husserl 24.5.1891' in *Philosophical and Mathematical Correspondence*, Brian McGuinness, ed. (Oxford: Blackwell 1980), 61-4, and reprinted in part as 'Letter to Husserl 24.5.1891' in *The Frege Reader*, 149-50. See also 'Über Begriff und Gegenstand,' *Vierteljahrsschrift für wissenschaftliche Philosophie* 16 (1892) 192-205, at 198; trans. by Max Black as 'On Concept and Object' in *Translations from the Philosophical Writings*, 42-55, and reprinted in *The Frege Reader*,

Second, in 'Function and Concept,' after citing his earlier criticism of formalism (in *The Foundations of Arithmetic*<sup>13</sup> and 'Formal Theories of Arithmetic'), Frege says:

Now what is the content, the referent of " $2 \cdot 2^3 + 2$ "? The same as of " $18$ " or " $3 \cdot 6$ ". What is expressed in the equation " $2 \cdot 2^3 + 2 = 18$ " is that the referent of the right-hand complex of signs is the same as that of the left-hand one.<sup>14</sup> ('Function and Concept,' 3)

According to Frege, ' $2 \cdot 2^3 + 2 = 18$ ' thus expresses that ' $2 \cdot 2^3 + 2$ ' and ' $18$ ' refer to the same object and, hence, he is again assuming the name view of identity statements.

Now, Frege's assertion of the name view in 'Formal Theories of Arithmetic' is perfectly consistent with the standard interpretation. After all, it was published seven years before 'Sense and Reference'; and the standard interpretation has Frege *changing his mind* about the name view in 'Sense and Reference.' And, since 'Function and Concept' comes the year before 'Sense and Reference,' it might seem that Frege's assertion of the name view there likewise does not tell against the standard interpretation. But with 'Function and Concept,' things are not as simple as they first appear. Between *The Foundations of Arithmetic* (1884) and the

181-93. And see *Grundgesetze der Arithmetik: begriffsschriftlich abgeleitet* [*The Basic Laws of Arithmetic*] vol. 1 (Jena: Pohle 1893), X and §5, 9 n.T; Preface, Introduction, and §§1-7, 26-9, and 32-3 trans. by Michael Beaney in *The Frege Reader*, 194-223. But Frege makes no such claim about the content of names that are not sentences. And although in a number of places he uses 'the content of a name (that is not a sentence)' for the referent of a name, nowhere does he use it for the sense of a name. For example, the statement of the name view in the *Begriffsschrift* (§8, 14) uses 'content' ('*Inhalt*') rather than 'referent' ('*Bedeutung*'). Later in that same section, Frege says of two names that refer to the same point: 'The name B has therefore in this case the same content as the name A.' And, in a passage from 'Function and Concept' quoted just below in the text, he uses 'content' and 'referent' interchangeably when speaking of names that are not sentences. See *Function und Begriff: Vortrag, gehalten in der Sitzung vom 9. Januar 1891 der Jenaischen Gesellschaft für Medizin und Naturwissenschaft* (Jena: Pohle 1891), 3; trans. by Peter T. Geach as 'Function and Concept' in *Translations from the Philosophical Writings*, 21-41, and reprinted in *The Frege Reader*, 130-48.

13 *Die Grundlagen der Arithmetik: eine logisch-mathematische Untersuchung über den Begriff der Zahl* (Breslau: Koebner 1884); trans. by J.L. Austin as *The Foundations of Arithmetic: A Logico-Mathematical Enquiry into the Concept of Number*, 2nd ed. (Evanston, IL: Northwestern University Press 1980).

14 We translate 'die Bedeutung der rechtsstehenden Zeichenverbindung dieselbe sei wie die der linksstehenden' as 'the referent of the right-hand complex of signs is the same as that of the left-hand one' rather than 'the right-hand complex of signs has the same referent as the left-hand one.'

first volume of *The Basic Laws of Arithmetic* (1893), Frege wrote 'Function and Concept,' 'Sense and Reference,' and 'On Concept and Object' together to develop his views of, among other things, truth-values and extensions of concepts. 'Function and Concept' and 'Sense and Reference' can thus be read as part of a single three-part work rather than as two papers written separately. And, given this, it would be at least a bit odd for Frege to publish the former even though it commits him to a view that he already rejects.

Although it makes Frege's discussion in 'Function and Concept' somewhat odd, we don't mean to suggest that this establishes that the standard interpretation of 'Sense and Reference' is incorrect. Even though the two papers were written together, Frege did publish 'Function and Concept' first; and there might be a perfectly good explanation for why he published it even though it commits him to a view he no longer held. But what does give us better reason to question the standard interpretation is that after 'Sense and Reference' Frege continues to assume the name view in criticisms of formalism.

In 1903, Frege published the second volume of *The Basic Laws of Arithmetic*, where he says:

We ourselves use the equality sign to express that the referent of the group of signs on the left-hand side coincides with the referent of the group of signs on the right. This is not applicable here [according to formalism], since a referent is not given. But we do not know what else the equality sign is to express. In any case, the signs on the left and on the right must then refer to something.<sup>15</sup>

In this passage, Frege not only says that an identity statement expresses that the names that flank the identity sign have the same referent; he also says he doesn't know what else the identity sign could express. Now, of course, Frege *does* know that the object view is a possible alternative account of the identity sign, since he discusses it in the opening paragraph of 'Sense and Reference.' Hence his remarks in the passage don't just establish that he holds the name view; they also establish that he takes the alternative object view to be completely untenable. And, later in the second volume of *The Basic Laws*, Frege praises Dedekind's non-formalist view of mathematics in *Continuity and Irrational Numbers*, from 1892. Frege says:

He says there in §1, p. 6: "To express the fact that the signs a and b refer to<sup>16</sup> one and the same rational number, one writes  $a = b$  or  $b = a$ ".

Here the sharpness of the distinction between sign and what it refers to is welcome and noteworthy, as is the conception of the equality sign, which exactly agrees with our own. (*The Basic Laws of Arithmetic*, vol. 2, §138, 140)

Notice that Dedekind's view of the identity sign just is the name view and that Frege says that Dedekind's view *exactly agrees with his own*. So, eleven years after 'Sense and Reference,' Frege twice asserts the name view in criticizing formalism and, on one of these occasions, his discussion strongly suggests that he takes the object view to be completely untenable. Moreover, these are not the only places after 'Sense and Reference' where Frege assumes the name view.

Frege also does so elsewhere in the second volume of *The Basic Laws*, in a section that has nothing to do with formalism. Frege considers and rejects the stipulation that mathematical expressions such as [the sum of  $\alpha$  and  $\beta$ ] have a referent only when  $\alpha$  and  $\beta$  refer to numbers. If this stipulation were correct, then, for example, 'the sum of the moon and the moon' would fail to have a referent and hence 'the sum of the moon and the moon is 1' and its negation, 'the sum of the moon and the moon is not 1,' would *both* fail to be either true or false. Frege rejects this idea in the following passage:

If our sentence "the sum of the Moon and the Moon is not one" were a scientific one, then it would state<sup>17</sup> that the words "the sum of the Moon and the Moon" and the word "one" did not coincide in reference; but with the stipulation suggested above, the former words would not have any referent; accordingly we could not truly state either that their referent did coincide with the referent of the word "one" or that it did not coincide with it. (*The Basic Laws of Arithmetic*, vol. 2, §64, 76)

In the first clause, Frege assumes that *the negation of* [ $\alpha = \beta$ ] states that  $\alpha$  and  $\beta$  do not co-refer, which is of course the negation of what [ $\alpha = \beta$ ] states on the name view. So Frege is once again assuming that the name view is the correct account of the identity sign.

So far, we have seen that Frege assumes the name view in 'Function and Concept,' which, although it was published one year before 'Sense and Reference,' was written at the same time as that work. We have also seen that Frege asserts the name view *three times after* 'Sense and Reference.'

15 *Grundgesetze der Arithmetik: begriffsschriftlich abgeleitet* [*The Basic Laws of Arithmetic*] vol. 2 (Jena: Pohle 1903), §105, 113. §§56-67, 138-47 trans. by Michael Beaney in *The Frege Reader*, 258-89. §§86-137 trans. by Max Black as 'Frege Against the Formalists' in *Translations from the Philosophical Writings*, 162-213.

16 We translate 'bedeuten' as 'refer to' rather than 'stand for' (similarly for 'bedeutet'). See note 2.

17 We translate 'besagte' here as 'state' (rather than 'assert'), as 'besagt' is translated in 'Formal Theories of Arithmetic' (101).

And, contrary to what we would expect if he rejects the name view in 'Sense and Reference,' nowhere does Frege indicate that his continued assertion of the name view is in any tension with anything he says in 'Sense and Reference.'

The passages quoted above should certainly raise some concern about the standard interpretation. And it's worth noting that the only way to read them as consistent with it is to suppose (as Bernhardt does) that, when Frege says, for example, 'What is expressed in the equation " $2 \cdot 2^3 + 2 = 18$ " is that the referent of the right-hand complex of signs is the same as that of the left-hand one,' all he means to be saying is that the equation is true if and only if the signs co-refer.<sup>18</sup> A proponent of the standard interpretation has to say that Frege is not making any claim at all about what thought the statement expresses, and likewise for the other passages quoted above. Now it is, of course, possible that Frege's choice of words in the passages quoted above — particularly 'state' ('besagt', 'besagte') and 'express' ('ausgedrückt', 'drücken') — is merely unfortunate and, hence, that these passages give no reason to be suspicious of the standard interpretation. But there's at least a *prima facie* puzzle as to why Frege would have made these unfortunate choices.

First, if he merely wanted to give truth-conditions for identity statements, he needn't have given them by using a view of identity that he rejects — that is to say, he could have said, for example, that ' $a = b$ ' is true if and only if *a and b are identical*, or if and only if *a and b are the same object*, without mentioning the names 'a' and 'b' at all. Indeed, this is exactly what he does in the first volume of *The Basic Laws*, where he says that " $\Gamma = \Delta$ " refers to the True if  $\Gamma$  is the same as  $\Delta$ ; in all other cases it refers to the False' (§7, 11). But second, and more importantly, if he merely wanted to give truth-conditions for identity statements, then why say that ' $2 \cdot 2^3 + 2 = 18$ ' expresses that the names co-refer? If the standard interpretation is right, Frege is using a claim that he has explicitly rejected (namely, the name view) but, so to speak, meaning something entirely different when it would have been just as easy for him to say what he meant. And it isn't just that Frege simply talks about what particular identity statements state or express; in the third passage he explicitly makes a claim about *what the identity sign expresses*, and in the fourth he explicitly makes a claim about *what his conception of the identity sign is*. These are, to put it mildly, enthusiastic ways of making a claim about mere truth-conditions; and, if that's what Frege is doing, he has gotten very carried away.

18 Stephen Bernhardt, 'Frege on Identity,' *Journal of Critical Analysis* 8.4 (1980) 57-65, at 61-4.

We think the above passages cast serious doubt on the idea that Frege rejected a metalinguistic account of identity in 'Sense and Reference.' However, there is one other place where Frege asserts the name view, and it is particularly damning for the standard interpretation of 'Sense and Reference.' In a letter to Giuseppe Peano, Frege says:

If we use the expression "the sun" in a sentence, we intend<sup>19</sup> by it a heavenly body which is in outer space and which is known to have mass. But the sense of the word "sun" is not somewhere in space, nor does it have mass. Thus if I write: " $5 + 2 = 4 + 3$ ", I do not intend by it that " $5 + 2$ " and " $4 + 3$ " have the same sense, but that they have the same referent, that they designate the same number.<sup>20</sup>

Here, Frege talks about what *he intends* by an identity statement; so, once again, it's more than a bit strained to claim he doesn't mean to be saying anything about the thought it expresses. And, once again, Frege's assertion occurs after the publication of 'Sense and Reference.' Although the letter is not dated, we know it was written sometime after the first volume of *The Basic Laws* was published in 1893, because Frege cites that work at the end of the letter. But there is another reason that the letter cannot have been written before 'Sense and Reference.' In a note to the last sentence of the passage quoted, Frege cites his discussion of sense and reference in '*Zeitschrift für Philosophie und philosophische Kritik*, vol. 100'; and that paper in volume 100 is none other than 'Sense and Reference.' This note conclusively shows that the letter to Peano comes after 'Sense and Reference,' but it shows much more than that; it also shows that *in a single sentence Frege both assumes that the name view of identity statements is correct and cites 'Sense and Reference.'* And, as with the other passages quoted in this section, Frege does not say that he has once again changed his mind; nor does he mention that anything odd might be going on. Hence the letter to Peano alone makes it almost out of the question that Frege is rejecting the name view in the opening paragraph of 'Sense and Reference.'<sup>21</sup>

19 We translate 'meine' and 'meinen' as 'intend' rather than 'mean,' since 'bedeuten' (which we translate as 'refer to') is sometimes translated as 'mean.' See note X. Z

20 'Frege an Peano ohne Datum,' *Wissenschaftlicher Briefwechsel*, 194-8, at 197; trans. by Hans Kaal as 'Frege to Peano undated,' in *Philosophical and Mathematical Correspondence*, 125-9.

21 On the name view,  $\lceil \alpha = \beta \rceil$  expresses the thought that  $\alpha$  and  $\beta$  co-refer. However, there is a closely related view — which we'll call the *hybrid view* — on which  $\lceil \alpha = \beta \rceil$  expresses the thought that the referent of  $\alpha =$  the referent of  $\beta$ . And, though the distinction is subtle, the thoughts  
(i) that  $\alpha$  and  $\beta$  co-refer

### III Frege's Prior Assertion of the Objections

There is no doubt that Frege raises an objection to the name view in the opening paragraph of 'Sense and Reference.' But we claim that, contrary to the standard interpretation, Frege does not think that the objection is a reason to abandon the view; rather, as we will see, he thinks it shows that one needs to be careful about how the view is formulated. As we have already seen, Frege asserts the name view at least four times after 'Sense and Reference.' This alone suggests that the standard interpretation cannot be correct, but there is more to be said against it. For 'Sense and Reference' is not the first place where Frege raises an objection to the name view; he raises the very same objection when he first states the

(ii) that the referent of  $\alpha$  = the referent of  $\beta$   
are in fact distinct. The thought (i) contains modes of presentation that determine the names  $\alpha$  and  $\beta$ , whereas the thought (ii) contains modes of presentation (given by 'the referent of  $\alpha$ ' and 'the referent of  $\beta$ ') that determine the objects  $a$  and  $b$ . So, like the object view, the hybrid view implies that identity is strictly speaking a relation between objects; however, it's important not to let this area of agreement obscure the fact that the views differ radically about the thoughts expressed by identity statements. For example, on the object view, 'Hesperus = Phosphorus' expresses the thought that Hesperus is Phosphorus; whereas, on the hybrid view, 'Hesperus = Phosphorus' expresses the metalinguistic thought (of type (ii)) that the referent of 'Hesperus' = the referent of 'Phosphorus.' In this respect, the hybrid view is much closer to the name view, on which 'Hesperus = Phosphorus' expresses the metalinguistic thought (of type (i)) that 'Hesperus' and 'Phosphorus' co-refer. Indeed, these metalinguistic thoughts are so closely related that Frege appears not to distinguish them. In three of the six passages cited in this section (i.e., the first, fourth, and sixth passages), what he says suggests that identity statements express thoughts of type (i), whereas in the other three passages (i.e., the second, third, and fifth ones) what he says suggests that identity statements express thoughts of type (ii) (given that, as he says in a footnote to 'Sense and Reference' (25 n.A), he is using 'coincides with' (*fallen zusammen*) and '=' interchangeably). We interpret Frege as holding the name view rather than the hybrid view because it fits better with the central texts: in Section 8 of the *Begriffsschrift*, Frege says that an identity statement expresses that the names 'have the same content' (*denselben Inhalt haben*) and thus identity is a relation between names (§8, 14); and, in the opening paragraph of 'Sense and Reference,' Frege (26) says that, on his view from the *Begriffsschrift*, an identity statement expresses that the names 'refer to the same thing' (*dasselbe bedeuten*) and that 'a relation between them [the names] would be asserted' (26). (On the interpretation of 'have the same content' (*denselben Inhalt haben*), see note 12 above.) But there might be other reasons to attribute the hybrid view rather than the name view to Frege. Given how subtle the distinction between the two views is, it would take another paper to settle the issue. However, whether Frege holds the name view or the hybrid view, the standard interpretation wrongly supposes that, in 'Sense and Reference,' Frege rejects his earlier view that identity statements express metalinguistic thoughts.

view in the *Begriffsschrift*.<sup>22</sup> And what he says there makes it difficult to see how he could be rejecting the name view in 'Sense and Reference.'

Recall that, in the opening paragraph of 'Sense and Reference,' Frege first raises an objection to the object view and then says:

What we apparently want to state by  $a = b$  is that the signs or names "a" and "b" refer to the same thing, so that those signs themselves would be under discussion; a relation between them would be asserted. But this relation would hold between the names or signs only in so far as they named or designated something. It would be mediated by the connection of each of the two signs with the same designated thing. But this is arbitrary. Nobody can be forbidden to use any arbitrarily producible event or object as a sign for something. In that case the sentence  $a = b$  would no longer touch the subject matter, but only its mode of designation; we would express no proper knowledge by its means. But in many cases this is just what we want to do. ('On Sense and Reference,' 26)

So his objection to the object view leads Frege, in the beginning of the passage, to say that the name view is apparently true; but, in the rest of the passage, he goes on to raise some kind of objection to the name view. This part of the passage is not easy to interpret, but one of Frege's worries about the name view seems to be that it gets the subject matter of identity statements wrong. On the name view, if you were to assert 'Hesperus is Phosphorus,' for example, you would be asserting that the names 'Hesperus' and 'Phosphorus' co-refer; those signs themselves would be under

22 There is one aspect of the name view that Frege mentions in the *Begriffsschrift* that is not mentioned in 'Sense and Reference': namely, that when a name flanks the identity sign it does not refer to the object it customarily refers to; rather, it refers to itself (§8, 13-14). This raises the question of why Frege does not mention this aspect of the view in 'Sense and Reference.' One answer is that, as we will see in Section VIII, Frege is not particularly concerned with identity in 'Sense and Reference.' But another answer is that, even in the *Begriffsschrift*, Frege is not so concerned with this aspect of the name view. He mentions it only in two, consecutive sentences; and, in the very first sentence where he mentions it, he proceeds to ignore it. For in that sentence he goes on to say that an identity statement expresses that the names refer to the same object; and they obviously don't refer to the same object if they refer to themselves (and aren't the same name).

As we will see in this section, although Frege does have an inchoate version of the sense-reference distinction in the *Begriffsschrift*, it is worth noting that, once the role of the sense-reference distinction is more fully worked out, the distinction seems to justify Frege's claim in the *Begriffsschrift* that a name that flanks the identity sign doesn't have its customary referent. After all, given the compositionality of reference, the referent of an expression is what's relevant to determining the truth-value (i.e., the referent) of a sentence it occurs in; and, on the name view, what's relevant to the truth-value of an identity statement are the names that flank the identity sign, since, on that view, an identity statement expresses the thought that those names co-refer.

discussion.' But, since 'nobody can be forbidden to use any arbitrarily producible event or object as a sign for something,' your assertion, to quote David Kaplan, turns out on the name view to be about 'what one might call an accident of human culture' rather than astronomy ('Words,' 118). So, as Frege goes on to say, on the name view 'a = b would no longer touch the subject matter.' Hence one of his worries is that the view implies, incorrectly, that all identity statements are about names.

But this does not explain Frege's remark that if the name view were correct 'we would express no proper knowledge' by means of an identity statement. After all, that two distinct names co-refer can surely count as *knowledge*, so wherein lies its impropriety? We think that Frege's worry that, on the name view, identity statements don't express proper knowledge is related to his first worry: namely, that they turn out to be about the wrong subject matter. That two names co-refer is not merely an accident of human culture; it is also a rather *uninteresting* accident of human culture. So the name view implies, incorrectly, that all identity statements express uninteresting information.

Although there has been much philosophical discussion of Frege's objections to the name view in 'Sense and Reference,' it has not been widely recognized that he considers the very same objections when he first states the name view in the *Begriffsschrift*. Immediately after stating the view there, he says:

This makes it appear at first as if it were here a matter of what pertains to the expression alone, not to the thought, and as if there were no need at all for different symbols for the same content and hence for a symbol for identity of content. (*Begriffsschrift*, §8, 14; original emphases)

The first part of the passage exactly expresses Frege's first objection to the name view in 'Sense and Reference': namely, that the name view gets the thought expressed by (or the subject matter of) identity statements wrong. The second part of the passage is more difficult to interpret: in what way does the name view imply or suggest that there is no need for different names for the same object and, hence, no need for the identity sign itself?

The question becomes easier to answer when we consider the second objection to the name view in 'Sense and Reference': namely, that identity statements turn out to be uninteresting. Suppose that, in the second part of the passage from the *Begriffsschrift*, Frege is thinking the same thing. Then it is not so hard to see how he could wonder why we would need different names for the same object at all if the view is correct. After all, if the identity sign is used only to express that two distinct names have the same referent, and if such facts are entirely uninteresting facts about our linguistic conventions, then why bother to have such conventions at all

— that is to say, why bother to have distinct names for the same object? And if there is no need to have distinct names for the same object, then there is no need for an identity sign on the name view; for on that view the identity sign is just used to express that two names co-refer. So Frege seems to be arguing that, if identity statements expressed uninteresting information, then there would be no reason to have two distinct names that refer to the same object, and, hence, that there would be no need for the identity sign itself. But notice that the argument doesn't really require mentioning that, on the name view, having two distinct names for the same object would be superfluous. For that is supposed to follow from the claim that, on the name view, identity statements turn out to be uninteresting. So the argument can be more simply presented: if the identity sign is never used to say anything of interest, then it would appear to be an utterly useless piece of language and, hence, one that could be dispensed with altogether. Indeed, it's hard to see what other reason there could be for dispensing with the identity sign.<sup>23</sup>

The upshot is that, when he formulates the name view in the *Begriffsschrift*, Frege is already aware of both of the objections to the name view in 'Sense and Reference.' And importantly, although Frege replies to the objections in the *Begriffsschrift*, he does not reply to them by abandoning the name view. By itself, none of this refutes the standard interpretation. After all, for all that has been said it is still possible that, by 1892, Frege had come to think that his previous reply to the objections was inadequate and indeed that these objections refuted the name view. So, the story might go, he repeats the objections in 'Sense and Reference,' but this time rejects the name view. Notice, however, that the story is already a little fishy. If it were true, then we would expect Frege to mention why he no longer took his reply in the *Begriffsschrift* to be compelling. But neither in 'Sense and Reference' nor anywhere else does Frege ever say that there is anything wrong with his reply in the *Begriffsschrift*. So, on the standard interpretation, Frege's discussion in 'Sense and Reference' constitutes a rather serious change of mind, which he fails to make any note of. And the story becomes even fishier when we consider Frege's reply to the objections in the *Begriffsschrift*. For what he says there, where he certainly is not rejecting the name view, is *essentially the same as what he says in 'Sense and Reference.'*

23 The only other reason for thinking that there would be no need for the identity sign on the name view is that, instead of saying [ $\alpha = \beta$ ], one could always say that  $\alpha$  and  $\beta$  co-refer. However, this can't be what Frege is thinking, because there is no way around this objection. Any analysis of identity in terms of something else will make the identity sign superfluous in this sense.



Frege's reply in the *Begriffsschrift* runs as follows:

To show that this appearance is deceptive, I take the following example from geometry. On the circumference of a circle lies a fixed point A, around which a straight line rotates. When the latter forms a diameter, let us call the point at the opposite end to A the point B. Let us then call, more generally, the point of intersection of the line and the circumference at any given moment the point B, resulting from the rule that continuous changes in position of the line always correspond to continuous changes in position of B. The name B therefore refers to<sup>24</sup> something indeterminate so long as the relevant position of the line is not given. It can now be asked: what point is yielded when the line is perpendicular to the diameter? The answer will be: the point A. The name B has therefore in this case the same content as the name A; and yet just one name could not have been used from the beginning, since the justification for doing so is only provided by this answer. The same point is determined in two ways:

- (1) immediately through intuition,
- (2) as the point B when the line is perpendicular to the diameter.

To each of these two modes of determination there corresponds a separate name. The need for a symbol for identity of content thus rests on the following: the same content can be fully determined in different ways; but that, in a particular case, *the same content* is actually given by *two modes of determination* is the content of a judgment. Before this judgment can be made, two different names corresponding to the two modes of determination must be provided for that [thing] that is thereby determined. But the judgment requires for its expression a symbol for identity of content to combine the two names. It follows from this that different names for the same content are not always merely a trivial matter of formulation, but touch the very heart of the subject matter<sup>25</sup> if they are connected with different modes of determination. (*Begriffsschrift*, §8, 14-15; original emphases)

Readers familiar with the opening paragraph of 'Sense and Reference' should be struck by the similarities between this passage and what Frege says after raising the objections to the name view in 'Sense and Reference.' As in 'Sense and Reference,' Frege points out that names have modes of determination (or modes of presentation) as well as referents; and, again as in 'Sense and Reference,' he uses a geometrical example to illustrate what modes of determination are.<sup>26</sup> Of course, it is possible that Frege has changed his mind about the role that modes of determination play in replying to the objections: it is possible that, in the *Begriffsschrift*, Frege thinks that modes of determination can be used to defend the name

view, while in 'Sense and Reference' he has come to think that the view is indefensible but that they can be used to defend the object view instead. However, given the similarities between the passages, we would expect Frege to draw at least *some* attention to the radically different roles these apparently similar passages are playing. Frege's silence about his alleged change of mind is even more puzzling given that he explicitly says in 'Sense and Reference' that he assumed the name view in the *Begriffsschrift*. For given the similarities between the objections to the name view and the reply in the *Begriffsschrift*, on the one hand, and the objections to the name view and what Frege says afterwards in 'Sense and Reference,' on the other, the reader who consults the *Begriffsschrift* after reading 'Sense and Reference' is apt to be puzzled as to the role that modes of presentation are playing in 'Sense and Reference.' So, on the standard interpretation, the apparent similarity between the *Begriffsschrift* reply and the 'Sense and Reference' reply masks the fact that they are entirely different; and, moreover, Frege makes no attempt to forestall the reader's confusion by noting this difference.<sup>27</sup>

27 On the standard interpretation there is another place where Frege appears to be oddly silent about a change in his view. As we pointed out in Section II, in the second volume of *The Basic Laws*, eleven years after 'Sense and Reference,' Frege agrees with Dedekind that the name view is the correct account of the identity sign. But immediately after expressing his agreement with Dedekind, Frege responds to an objection of Thomae's concerning identity by citing the sense-reference distinction. Hence, if the standard interpretation is correct, Frege must again have changed his mind about which view the distinction makes defensible and again failed to make any note of the change. With this said, there are difficulties in interpreting Frege's criticism of Thomae. Thomae's objection in *Elementary Function Theory* from 1898 is that 'if equality or the equality sign = were only to stand for identity, then we would be left with trivial knowledge, or if one prefers, the conceptual necessity a is a (a = a)' (quoted in *The Basic Laws* vol. 2, §138, 140). So Thomae's objection seems to be the objection to the object view: namely, that if  $\lceil \alpha = \beta \rceil$  expresses that a and b are the same object, then it would seem to express the same thing as  $\lceil \alpha = \alpha \rceil$ . Since Frege replies to Thomae by citing the distinction between sense and reference, it might seem that Frege is using that distinction to defend the object view. But Frege asserts neither the object view nor the name view in his reply to Thomae. Frege says that senses determine cognitive value and that the senses of the two identity statements can differ, but he does not say *how* they differ. More importantly, as we pointed out at the beginning of this note, the quotation from Thomae and Frege's reply occur *immediately after, and continue the discussion started in,* the citation from Dedekind and Frege's claim that Dedekind's view 'exactly agrees with our own.' It is because of these interpretive difficulties that we have not discussed the criticism of Thomae in the text. But however these difficulties are resolved, it is clear that Frege appeals to senses (or modes of presentation) to reply to an objection to a view about identity; and it is clear that Frege is at the same time assuming that the name view is true.

24 We translate 'bedeutet' as 'refers to' rather than 'denotes.' See note 2.

25 On the translation of 'Sache selbst' as 'subject matter' (rather than 'matter'), see note 4 above.

26 We use 'mode of determination' and 'mode of presentation' interchangeably. As far as we can tell, there is no reason to think that the change of terminology amounts to anything.

So what Frege says in 'Sense and Reference' after raising his objections to the name view is remarkably similar to his reply to those objections in the *Begriffsschrift*. But the similarity doesn't stop at the fact that both involve the distinction between modes of determination and referents; a close reading of both passages reveals that the replies in them are more or less the same. To see this, we first need to see how Frege's reply in the *Begriffsschrift* is meant to address the objections.

#### IV The *Begriffsschrift* Reply

Recall that, in the *Begriffsschrift*, Frege raises two related objections to the name view: first, it gets the subject matter of most identity statements wrong; and, second, it makes it hard to see why we need a sign for identity. In his reply (quoted in the previous section) to the objections, Frege begins by using an example from geometry in which different names that have the same referent are associated with different modes of determination of that referent. He continues:

The need for a symbol for identity of content thus rests on the following: the same content can be fully determined in different ways; but that, in a particular case, the same content is actually given by two modes of determination is the content of a judgment. Before this judgment can be made, two different names corresponding to the two modes of determination must be provided for that [thing] that is thereby determined. But the judgment requires for its expression a symbol for identity of content to combine the two names. (*Begriffsschrift*, §8, 14-15; original emphases)

Two minor problems arise in interpreting this passage. First, although Frege is replying to an objection to his view of identity statements, he talks about the content of identity judgments. But Frege must be assuming that the content of a judgment just is the content of a statement used to express the judgment. So his remarks about identity judgments can be read as remarks about identity statements and, hence, are responsive to the objections. Second, in the first sentence Frege talks sometimes of the content of a sentence, sometimes of the content of a name. As we pointed out above (in note 12), Frege later says that, at the time of the *Begriffsschrift*, the content of a sentence is sometimes its truth-value, sometimes the thought it expresses. (See 'Frege to Husserl 24.5.1891,' 97; 'Concept and Object,' 198; and *The Basic Laws* vol. 1, X and §5, 9 n.T.) But the content of a name (that is not a sentence) is always its referent, never its sense. In the first sentence, the content that is determined by modes of determination (that 'can be fully determined in different ways' and 'is actually given by two modes of determination') is clearly the referent of a name: that is, it's an object. The content of an identity judgment, then, is that the same object is determined by two different modes of determina-

tion. So the content of an identity statement is also that the same object is determined by two different modes of determination. Here, the content of an identity statement cannot be a truth-value, since that the same object is determined by two different modes of determination is not a truth-value. So the content of an identity statement must be the thought it expresses. What Frege says in the first sentence quoted above is thus that  $\lceil \alpha = \beta \rceil$  expresses the thought that the modes of determination associated with  $\alpha$  and  $\beta$  determine the same object. This is Frege's answer to the second objection to the name view: we need an identity sign in our language to express the thought that the different modes of determination associated with the relevant names determine the same object.

Now, although Frege's claim about why we need the identity sign seems similar to the name view, it is actually a different view. On the name view, an identity statement expresses that the names that flank the identity sign co-refer. But the claim that Frege makes in his explanation of why we need an identity sign is that an identity statement expresses that the modes of determination associated with those names determine the same object. Now it might be thought that Frege is rejecting the name view in favor of the mode of determination view of identity statements, the view that an identity statement expresses that the modes of determination associated with the relevant names determine the same object.<sup>28</sup> But, however tempting it might be, this thought cannot be right. In the opening paragraph of 'Sense and Reference,' Frege says that he assumed the name view in the *Begriffsschrift* and makes no mention of the mode of determination view. And, four sentences after the passage quoted above, Frege ends his reply in the *Begriffsschrift* to the objections by again asserting the name view. He says:

$\lceil \text{—} (A \equiv B) \rceil$

is therefore to mean: the symbol A and the symbol B have the same conceptual content, so that A can always be replaced by B and vice versa.<sup>29</sup> (*Begriffsschrift*, §8, 15; original emphases)

28 Notice that, like the name view and the object view, the mode of determination view has consequences for what the referent of a name that flanks the identity sign is. The object view implies that the referent of a name that flanks the identity sign is its customary referent and, as pointed out in note 22, the name view implies that the referent of a name that flanks the identity sign is the name itself. On the mode of determination view, the referent of a name that flanks the identity sign will be the name's customary mode of determination; for, according to the mode of determination view, an identity statement expresses the thought that the modes of determination associated with the names determine the same object.

29  $\lceil \text{—} \rceil$  is the assertion sign. And, as Frege makes clear in the first volume of *The Basic*

So Frege is certainly not replying to the objections by rejecting the name view in favor of the mode of determination view.

Frege's official view of identity statements in the *Begriffsschrift* is the name view. Hence the only explanation for his assertion of the mode of determination view is that, at this stage, he is running the two views together; he simply does not see that the name view and the mode of determination view are different views of identity statements. This no doubt seems like an odd conflation for Frege to make, but in the next section we will see that, given what he says about the relationship between names and their modes of presentation in 'Sense and Reference,' it is far less odd than it first appears.

However, before turning to this, we need to see how Frege takes himself to be answering the objections in the *Begriffsschrift*. As we have seen, Frege claims that we need an identity sign in order to assert that two distinct modes of determination determine the same object. But recall that the reason Frege was worried about why we might need an identity sign is that the name view appears to have the consequence that identity statements turn out to be about language and, hence, turn out to be uninteresting. So his explanation of why we need the identity sign isn't adequate unless he can explain the following: why does the (alleged) fact that identity statements express that two modes of determination determine the same object imply that identity statements aren't just about language and, hence, can be interesting? Frege addresses this question immediately after his claim that we need an identity sign in order to state that two modes of determination determine the same object. He says:

It follows from this that different names for the same content are not always merely a trivial matter of formulation, but touch the very heart of the subject matter if they are connected with different modes of determination. (*Begriffsschrift*, §8, 15)

Frege has raised two objections to the name view: (i) it gets the subject matter of most identity statements wrong; and (ii) it implies, often incorrectly, that identity statements are trivial and uninteresting and, hence, that there is no need for an identity sign. We can thus read Frege as replying that, if the relevant names are associated with different

modes of determination, then (i) the name view does not get the subject matter of most identity statements wrong (identity statements are not merely a 'matter of formulation'); and (ii) the name view does not imply that identity statements are trivial and uninteresting (identity statements are not 'a trivial matter'). But this raises the question of why Frege thinks that, if  $\alpha$  and  $\beta$  are associated with different modes of presentation,  $\lceil \alpha = \beta \rceil$  has the right subject matter and is potentially interesting.

Given that Frege is running the name view and the mode of determination view together, it is not hard to see what the answer to this question is. On the mode of determination view,  $\lceil \alpha = \beta \rceil$  expresses that the mode of determination associated with  $\alpha$  and the mode of determination associated with  $\beta$  determine the same object. And, in the example that Frege discusses, the modes of determination themselves do in fact 'touch the very heart of the subject matter.' Recall that, in the geometrical example from the *Begriffsschrift*, the mode of determination associated with 'A' determines the point 'immediately through intuition,' whereas the mode of determination associated with 'B' is 'the point when the line is perpendicular to the diameter.' So, if 'A = B' expresses that these two modes of determination determine the same object, it's clear why someone might think that the identity statement does touch the heart of the subject matter. For, plausibly, if you know that these two modes of determination determine the same point, you will know that the point when the line is perpendicular to the diameter is the same as the point determined immediately through intuition. On the mode of determination view, it thus seems that knowledge that the identity statement is true implies genuine geometrical knowledge, because the modes of determination themselves concern the right subject matter.

Unfortunately, although it *seems* that genuine geometrical knowledge is implied, in fact it is not. You might know that the mode of determination associated with 'A' and the mode of determination associated with 'B' determine the same object without having any idea what either mode of determination is; and, if you do not have any idea what either mode of determination is, knowing that the two modes of determination determine the same object will not give you any geometrical knowledge at all. With this said, it is plausible that Frege is tacitly assuming that understanding a name requires knowing which mode of determination is associated with it.<sup>30</sup> And with this assumption in place, his account of

<sup>30</sup> Laws, he eventually replaces ' $\equiv$ ' with '=' (IX). The reader might be tempted to make something of Frege's use of 'conceptual content' rather than 'content'; however, in the preface to the *Begriffsschrift*, Frege makes it clear that the conceptual content of a name is just its referent. He says: 'I have called, in §3, that which solely mattered to me *conceptual content*' (IV; original emphases). And, with regard to names, the content that matters to Frege in the *Begriffsschrift* is the name's referent.

<sup>30</sup> Frege assumes a related version of this claim in 'Sense and Reference.' He says: 'The sense of a proper name is grasped by everybody who is sufficiently familiar with the language or totality of designations to which it belongs' rather than saying that the sense is grasped by anyone who merely understands *the name itself* (27).

why identity statements touch the very heart of the subject matter makes sense.

## V The Opening Paragraph of 'Sense and Reference'

On the standard interpretation of the opening paragraph of 'Sense and Reference,' Frege takes his objections to the name view to be decisive and hence rejects the name view in favor of the object view. As we have seen, the standard interpretation makes it difficult to make sense of many of Frege's other writings. To sum up:

- (i) In 'Function and Concept,' where he first introduces the distinction between sense and reference, Frege asserts the name view.
- (ii) In at least four places after 'Sense and Reference,' Frege asserts or otherwise commits himself to the name view; and, in one of those places, he cites 'Sense and Reference' in the same sentence.
- (iii) In the *Begriffsschrift*, where he first asserts the name view, Frege has already considered and replied to the objections to the view that occur in 'Sense and Reference.' But, in the *Begriffsschrift*, he certainly does *not* think that these objections provide a reason for rejecting the name view, and his response to them there is extremely similar to his response in 'Sense and Reference.'

These points amount to a mountain of circumstantial evidence that the standard interpretation of 'Sense and Reference' cannot be correct. However, we think that a careful examination of the opening paragraph of 'Sense and Reference' *itself* shows that Frege is not rejecting the name view there. Indeed, a careful examination of the opening paragraph shows that, as one would expect given their surface similarities and Frege's complete silence about any difference between them, the discussion there is remarkably similar to the discussion in the *Begriffsschrift*. Here, once again, is the first paragraph of 'Sense and Reference,' this time carved at the joints.

- I. Equality gives rise to challenging questions which are not altogether easy to answer. [Footnote to "Equality": I use this word in the sense of identity and understand "a = b" to have the sense of "a is the same as b" or "a and b coincide."] Is it a relation? A relation between objects, or between names or signs of objects?
- II. In my *Begriffsschrift* I assumed the latter. The reasons which seem to favor this are the following: a = a and a = b are obviously statements of differing cognitive value; a = a holds *a priori* and, according to Kant, is to be labeled analytic, while

statements of the form a = b often contain very valuable extensions of our knowledge and cannot always be established *a priori*. The discovery that the rising sun is not new every morning, but always the same, was one of the most fertile astronomical discoveries. Even today the reidentification of a small planet or comet is not always a matter of course. Now if we were to regard equality as a relation between that which the names "a" and "b" refer to, it would seem that a = b could not differ from a = a, i.e. provided that a = b is true. A relation would thereby be expressed of a thing to itself, and indeed one in which each thing stands to itself but to no other thing. What we apparently want to state by a = b is that the signs or names "a" and "b" refer to the same thing, so that those signs themselves would be under discussion; a relation between them would be asserted.

- III. But this relation would hold between the names or signs only in so far as they named or designated something. It would be mediated by the connection of each of the two signs with the same designated thing. But this is arbitrary. Nobody can be forbidden to use any arbitrarily producible event or object as a sign for something. In that case the sentence a = b would no longer touch the subject matter, but only its mode of designation; we would express no proper knowledge by its means. But in many cases this is just what we want to do.
- IV. If the sign "a" is distinguished from the sign "b" only as an object (here, by means of its shape), not as a sign (i.e. not by the manner in which it designates something), the cognitive value of a = a becomes essentially equal to that of a = b, provided that a = b is true. A difference can arise only if the difference between the signs corresponds to a difference in the mode of presentation of the thing designated. Let a, b, c be the lines connecting the vertices of a triangle with the midpoints of the opposite sides. The point of intersection of a and b is then the same as the point of intersection of b and c. So we have different designations for the same point, and these names ("point of intersection of a and b," "point of intersection of b and c") likewise indicate the mode of presentation; and hence the statement contains actual knowledge. ('On Sense and Reference,' 25-6; 25 n.A)

In part I, Frege introduces the two views of identity. In part II, he raises an objection to the object view and says that this objection was his reason for adopting the name view in the *Begriffsschrift*. In part III, he raises the objections to the name view. If the standard interpretation is correct, then it is in part IV that Frege rejects the name view in favor of the object view. However, a careful reading of the fourth part reveals that, rather than rejecting the name view, Frege is clarifying it.

The first thing to notice is that *nowhere in the first paragraph does Frege say that he is rejecting the name view; nor does he ever assert the object view.* But if Frege neither explicitly rejects the name view nor explicitly accepts the object view, what is it about the opening paragraph of 'Sense and Reference' that makes it seem that he is rejecting the name view? We think four features of the passage are responsible for this mistaken impression. Importantly, they are all suggestive rather than conclusive, and none occurs in the crucial fourth part. The first three are as follows:

- (i) At the beginning of the second part, rather than simply *asserting* the name view, Frege only says that he *assumed* it in the *Begriffsschrift*.
- (ii) In the middle of the second part, Frege says the objections to the object view *seem* to favor the name view, rather than simply saying that they *do* favor it.
- (iii) At the end of the second part, Frege says that we *apparently want* to say that identity is a relation between names, rather than saying that this is what *should* be said.

In each of these cases, Frege makes a weaker claim than we might expect if he still thinks that the name view is correct. But the qualified nature of Frege's remarks can be explained by the fact that he is about to raise objections to the name view in the third part. Moreover, although (i)-(iii) may give the impression that Frege is about to reject the name view, they certainly don't give sufficient reason to believe that he does. Whatever Frege's final view of identity is, it comes in the fourth part. Hence any interpretation of the opening paragraph must stand or fall by what it says goes on in this fourth part. However, before turning to the fourth part, we need to consider one more feature of the passage that gives the impression that Frege is rejecting the name view.

This last feature comes in the third part, where Frege raises an objection to the name view. After saying that we apparently want to say that the name view is correct, he says that the connection between names and objects is arbitrary, because 'nobody can be forbidden to use any arbitrarily producible event or object as a sign for something.' He then says that 'in that case the sentence  $a = b$  would no longer touch the subject matter' and 'we would express no proper knowledge by its means,' but that 'in many cases this is just what we want to do.' Frege in no way qualifies his claim that the name view has these consequences (for example, he does not say that the name view *seems* to imply that we would not express proper knowledge), nor does he qualify his claim that these consequences are unacceptable in many cases (for example, he does not say that it *seems* that we want to express proper knowledge in many cases). That Frege says, without qualification, that the name view has certain consequences and says, again without qualification, that these consequences are unacceptable suggests that he thinks that the objections to the name view cannot be met. This does provide some reason to think that Frege no longer accepts the name view. But what he says next is far from a rejection of the name view. In fact, the fourth part of the opening paragraph of 'Sense and Reference' must be read as a *reply* to the objections to the name view.

At the beginning of the fourth part, immediately after raising the objections to the name view, Frege says:

If the sign "a" is distinguished from the sign "b" only as an object (here, by means of its shape), not as a sign (i.e. not by the manner in which it designates something), the cognitive value of  $a = a$  becomes essentially equal to that of  $a = b$ , provided that  $a = b$  is true. A difference can arise only if the difference between the signs corresponds to a difference in the mode of presentation of the thing designated. ('On Sense and Reference,' 26)

Frege then goes on to give an example involving two distinct geometrical modes of presentation (the point of intersection of l and m, and the point of intersection of m and n) and says that, if two distinct names are associated with these different modes of presentation, the identity statement containing the names 'contains actual knowledge.'<sup>31</sup> He in no way says how the fact that the two names are associated with distinct modes of presentation implies that the identity statement contains actual knowledge and, hence, the example doesn't help to determine whether he's using modes of presentation to defend the name view (as in the *Begriffsschrift*) or whether he is now using them to defend the object view. So the question of which view he is defending in the opening paragraph of 'Sense and Reference' depends *entirely on the two sentences quoted above*.

The first point to be made is that, whichever objections Frege is replying to, his reply is incredibly brief. This fact alone raises more worries for the standard interpretation. For, if, as the standard interpretation supposes, Frege is changing his mind about which view modes of presentation can be used to defend, one would not only expect him to make some note of his change of mind; one would also expect him to elaborate on his reply here so that the reader can see that it differs from the *Begriffsschrift* reply. But the two sentences also raise more substantive worries for the standard interpretation.

The first sentence is a conditional: Frege says that, if the names (or signs) are distinguished as objects (that is, orthographically, by their shapes) rather than as names (that is, semantically, by the modes of presentation associated with them), then the cognitive values of  $[a = a]$  and  $[a = b]$  will turn out to be essentially the same. The second sentence is also a conditional, and it seems to merely restate the first sentence as a contrapositive: if the cognitive values of the two sentences are to be different, the names must be associated with distinct modes of presen-

31 In 'Sense and Reference,' Frege uses 'a,' 'b,' and 'c' as the names of the lines. Since he also uses ' $a = b$ ' as an identity statement, we've changed the names of the lines to avoid confusion.

tation. Now, whichever view Frege is defending here, he clearly thinks that he has already established that the cognitive values of the two sentences aren't (or at least needn't be) essentially the same — that is to say, he clearly thinks that he has established that the consequent of the first conditional is false. Hence we are obviously supposed to conclude by *modus tollens* that the names must be associated with distinct modes of presentation. But is he saying that the names must be associated with distinct modes of presentation if *the object view* is to avoid the consequence that the cognitive values of the two identity statements are essentially the same? Or is he saying that the names must be associated with distinct modes of presentation if *the name view* is to avoid this consequence? So the question of which view Frege is defending can be settled by figuring out whether he's talking about the name view or the object view in the first sentence. That is to say, it can be settled by figuring out which of the following is the correct interpretation of what Frege means in the first sentence.

#### Name View Interpretation

If the sign "a" is distinguished from the sign "b" only as an object (here, by means of its shape), not as a sign (i.e. not by the manner in which it designates something), then **on the name view** the cognitive value of  $a = a$  becomes essentially equal to that of  $a = b$ , provided that  $a = b$  is true.

#### Object View Interpretation

If the sign "a" is distinguished from the sign "b" only as an object (here, by means of its shape), not as a sign (i.e. not by the manner in which it designates something), then **on the object view** the cognitive value of  $a = a$  becomes essentially equal to that of  $a = b$ , provided that  $a = b$  is true.

If the former interpretation is correct, he is defending the name view; and, if the latter is correct, he is defending the object view. But there are two reasons to think that the latter interpretation isn't correct and, hence, that Frege is in fact defending the name view.

First, the object view interpretation makes a certain feature of the antecedent of the conditional mysterious. Recall that Frege's worry about the object view is that it implies that the two identity statements say *exactly* the same thing: each says of the same object that it is identical with itself. But, given this, why would Frege say that, if the names are individuated by their shapes rather than by their modes of presentation, the cognitive values of the sentences turn out to be only *essentially* the same? That is to say, why wouldn't he claim that they turn out to be *exactly* the same? However, if the name view interpretation is correct, Frege's insertion of the word 'essentially' ('wesentlich') is easily ex-

plained. The cognitive values of the identity statements are *essentially* the same, because, as we pointed out in Section III, on the name view they both express uninteresting metalinguistic thoughts; hence the difference between the thoughts expressed by the sentences isn't sufficient to explain their distinct cognitive values.<sup>32</sup>

Now, in the end, we wouldn't want to rest too much of our case on Frege's use of the word 'essentially.' And, fortunately, we don't have to. For, if the object view interpretation is correct, the first sentence of the quotation constitutes a fantastic non sequitur. In the two sentences immediately before, Frege continues his objection to the name view by saying it implies that 'we would express no proper knowledge' by an identity statement, even though there are many cases in which we do use identity statements to do just that. But, given this, it isn't at all tenable that he is, without comment, returning the discussion to the object view. On the standard interpretation, the first sentence of the passage and the sentence immediately preceding it have to be read as follows:

In that case, **on the name view**, the sentence  $a = b$  would no longer touch the subject matter, but only its mode of designation; we would express no proper knowledge by its means. But in many cases this is just what we want to do. If the sign "a" is distinguished from the sign "b" only as an object (here, by means of its shape), not as a sign (i.e. not by the manner in which it designates something), then **on the object view** the cognitive value of  $a = a$  becomes essentially equal to that of  $a = b$ , provided that  $a = b$ .

The standard interpretation is committed to this reading of these three sentences; and, while the non sequitur may not be numbingly gross, it could only come from a numbingly disorganized thinker. If Frege really were changing the subject in the way that the standard interpretation implies, he would certainly have made it clear that his discussion of the name view had come to an end and that the defects of the view that the discussion revealed force a reconsideration of the object view. But, of course, Frege does no such thing and, hence, the standard interpretation of the opening paragraph is simply wrong.

32 In the second sentence, Frege does say that a difference can arise only if the two names are associated with distinct modes of presentation. And it might be claimed that his talk of a *difference* rather than an *interesting* or *substantive* difference suggests that he's offering a defense of the object view. But the second sentence appears to be nothing more than a contrapositive of the first sentence; hence, given that he talks of *essential* sameness in the first sentence, it is natural to read 'a difference' in the second sentence as referring to a difference that wouldn't make the cognitive values essentially the same.

To sum up the results of this section:

- (i) The fourth part of the opening paragraph of 'Sense and Reference' is very brief — too brief to plausibly be read as a repudiation of his discussion in the *Begriffsschrift*.
- (ii) Frege's reply is meant to address the worry that the cognitive values of the two identity statements turn out to be *essentially* the same. This doesn't make sense if he is replying to the objections to the object view, but it makes perfect sense if he is replying to the objections to the name view.
- (iii) Frege's reply occurs immediately after his discussion of the objections to the name view and he gives no indication whatsoever that the focus of the discussion has changed. Hence, if he is defending the object view, he must have committed an incredible editorial blunder.

In Sections II-IV we presented evidence from texts other than 'Sense and Reference' that Frege never changes his mind about the name view. In this section, we think we've shown that a careful reading of the opening paragraph of 'Sense and Reference' itself shows that, as in the *Begriffsschrift*, Frege is using modes of presentation to defend the name view. Of course, if this is correct, then some account must be given of how Frege's claim that names are to be distinguished by their modes of presentation is supposed to rescue the name view from the objections in 'Sense and Reference.' We think that such an account can be given. Moreover, the account sheds light on the *Begriffsschrift* reply to the objections and shows — as one would expect given that Frege makes no comments at all about any differences between them — that there is a very close connection between the replies in the *Begriffsschrift* and 'Sense and Reference.'

## VI The 'Sense and Reference' Reply

Recall that Frege's *Begriffsschrift* reply to the objections conflates the view that an identity statement expresses a relation between *names* with the view that it expresses a relation between the *modes of determination associated with the names* — his reply conflates names with the modes of determination associated with them. Frege does not make this conflation in his 'Sense and Reference' reply to the objections, but what he winds up saying isn't, in the end, all that different. The 'Sense and Reference'

reply is, as it were, a way of putting the *Begriffsschrift* reply so that it avoids the conflation.

Recall that Frege's claim about the individuation conditions of signs occurs in the following sentence:

If the sign "a" is distinguished from the sign "b" only as an object (here, by means of its shape), not as a sign (i.e. not by the manner in which it designates something), the cognitive value of  $a = a$  becomes essentially equal to that of  $a = b$ , provided that  $a = b$  is true. ('On Sense and Reference,' 26)

The parenthetical remarks in the antecedent of the conditional indicate how signs are individuated as signs and how they are individuated as mere objects. These remarks are pretty compressed, but Frege seems to be saying that to individuate signs as *signs* is to individuate them by the modes of presentation associated with them, whereas to individuate them as *mere objects* is (in the case of written signs) to individuate them by their shapes. As stated in the last section, it's clear that Frege takes the consequent of the conditional to be false (the cognitive values of  $\lceil \alpha = \beta \rceil$  and  $\lceil \alpha = \alpha \rceil$  aren't essentially the same) and, hence, that he is claiming that signs must be individuated by their modes of presentation. Although this claim itself is a bit compressed, it isn't at all difficult to see what it amounts to.

The claim about how signs must be individuated certainly isn't a claim about what makes two *token* signs numerically identical; regardless of which two signs flank the identity sign, the two tokens are numerically distinct, since one occurs to the left of the identity sign, while the other occurs to the right. Hence the claim must concern the conditions under which two *token* signs are of the same *type*; and what Frege is saying is that the conditions depend upon which type is under discussion. The two tokens are of the same *object* type just in case they have the same shape, but they are of the same *sign* type just in case they are associated with the same mode of presentation. For example, consider the two words below:

Color

Colour

Are they of the same type? Well, if the question is whether they are of the same object type, the answer (according to Frege) is no: the first word contains five parts, while the second contains six, so their shapes are different. But if the question is whether they are of the same sign type, the answer (according to Frege) is yes: the signs are of exactly the same semantic type, so they are associated with the same mode of presentation. Or consider the sentences below:

Banks loan money.

Banks are river edges.

Is the first word of the sentence on the left of the same type as the first word of the sentence on the right? Again, this depends upon what type is under discussion. If the question is whether they are of the same object type, the answer (according to Frege) is yes: both have the same shape. But, if the question is whether they are of the same sign type, the answer (according to Frege) is no: the signs are of different semantic types (assuming the sentences are both true), so they are associated with different modes of presentation.<sup>33</sup>

33 We have interpreted Frege as saying that sameness of mode of presentation type is both necessary and sufficient for sameness of sign type. This implies that any two exact synonyms are of the same sign type; but it is certainly implausible that, for example, 'doctor' and 'physician' count as the same sign. Hence it might be thought that a more charitable interpretation would have Frege saying that the individuation conditions of signs include those of objects — that sameness of mode of presentation type and sameness of object type are *both* required for sameness of sign type. Now in the end nothing that we say depends upon accepting the interpretation in the text over this alternative. But we have interpreted Frege as we do for two reasons. First, he talks of distinguishing 'a' from 'b' as an object and 'not as a sign,' and this suggests that the individuation conditions of objects are *not* included in those of signs. And second, he talks of individuating 'a' from 'b' 'not as a sign (i.e., not by the manner in which it designates something)'; and his use of 'i.e.' suggests that to *not* individuate names as signs just is to *not* individuate them by their modes of presentation. Of course, neither reason is conclusive; but notice that the alternative interpretation suggested above is not really more charitable to Frege. Granted, if Frege holds that sameness of sign type requires sameness of mode of presentation type as well as sameness of object type, then 'doctor' and 'physician' turn out to be different signs on his view. But, implausibly, 'color' and 'colour' *also* turn out to be different signs; and, if we individuate spoken signs in part by their sounds, dialectical differences in spoken language (for example, the British and American pronunciations of 'schedule') will also imply that the relevant signs are different. The alternative interpretation simply trades one set of counterintuitive consequences for another and, hence, we stick with the one that we think best fits the text.

It should also be noted that Frege says: 'In the case of an actual proper name such as "Aristotle" opinions as to the sense may differ' (27 n.B); and this might suggest that he thinks that two token names can be of the same sign type even if they are associated with distinct modes of presentation. But it is plausible that, when Frege says that there can be differences in opinion as to the sense attached to a name, he is individuating names as objects rather than as signs. Recall that *names* can be individuated as objects. Moreover, at the end of the footnote, Frege says that 'such variations of sense may be tolerated,' but that they 'ought not to occur in a perfect language' (27 n.B). So, even if Frege is individuating 'Aristotle' as a name, it's very plausible that his claims about the way signs are individuated *qua* signs are meant to apply only to a perfect language and that what makes a language in which there are such variations imperfect is (at least in part) that the relevant signs in such a language aren't individuated correctly *qua* signs.

Finally, since any identity statement occurs in a single language, there is no reason to think that Frege is committed to the implausible claim that two token signs

As we suggested at the beginning of this section, if Frege does hold this view of the individuation conditions of sign types, his conflation of names with their modes of determination in the *Begriffsschrift* isn't so mysterious. For names are signs; and, if sameness of mode of determination is both necessary and sufficient for sameness of sign type, then *sameness of mode of determination amounts to the same thing as sameness of name*; the claim that the names that flank the identity sign are the same (that is, of the same sign type) amounts to the claim that the modes of determination associated with them are the same. Frege's puzzling conflation of the name view and the mode of determination view isn't so puzzling after all if he also held (perhaps implicitly), when he wrote the *Begriffsschrift*, that signs *qua* signs are individuated entirely by the modes of determination associated with them.<sup>34</sup> The *Begriffsschrift* reply and the 'Sense and Reference' reply turn out to be very similar. But, in 'Sense and Reference,' rather than conflating names with their modes of presentation, Frege is claiming that sameness of name *just amounts to* sameness of mode of presentation.

Moreover, the fact that Frege takes names *qua* signs to be entirely individuated by their semantic properties allows us to make sense of his 'Sense and Reference' reply in roughly the same way that we made sense of his parallel *Begriffsschrift* reply. Frege's examples of modes of presentation in 'Sense and Reference' are again geometrical: the same point is presented as the point of intersection of the distinct lines l and m, and as the point of intersection of the distinct lines m and n. Suppose that the name 'a' is associated with the first mode of presentation and the name 'b' is associated with the second. Given that what makes the names tokens of distinct sign types is that they are associated with distinct modes of presentation, it isn't implausible to suppose that anyone who knows that the names are tokens of distinct sign types will know which mode of presentation is associated with each of them. Now this doesn't

from distinct languages are of the same sign type if they are associated with the same mode of presentation. So the claim about sign individuation should probably be qualified so that it applies only to signs from the same language. For simplicity's sake, we ignore this complication in the text.

34 We do not mean to claim that Frege explicitly held this view when he wrote the *Begriffsschrift*. Although it is possible that he did explicitly hold it, his conflation of names with the modes of determination associated with them is explained if he was in some way merely implicitly inclined to think that linguistic items are individuated entirely by their semantic properties. Although this seems to be the only way to explain his conflation, the explanation doesn't require that his explicit thoughts on the subject were very clear.



yet explain how the name view avoids the objections that it fails to get the subject matter of identity statements right and that it implies that the knowledge they express is always trivial. But since Frege's geometrical examples of modes of presentation concern the subject matter of geometry, knowledge that two such modes of presentation determine the same object will yield significant geometrical knowledge. If you know that the distinct names 'a' and 'b' co-refer, and if what makes them distinct names is that one is associated with the mode of presentation given by 'the intersection of lines l and m' while the other is associated with that given by 'the intersection of lines m and n,' then plausibly you'll know that the intersection of l and m is the intersection of m and n. And this is significant geometrical knowledge.

On this interpretation, Frege's 'Sense and Reference' reply is very similar to his *Begriffsschrift* reply and — as we've continually emphasized above — we think this is as it should be. The similarities between the *Begriffsschrift* objections and reply, on the one hand, and the 'Sense and Reference' objections and reply, on the other, make it incredible that Frege should be making radically different claims without commenting on the difference. However, on the interpretation we've given, the 'Sense and Reference' reply, like the *Begriffsschrift* reply, doesn't really answer the objections. For, even if name types are entirely individuated by the modes of presentation associated with them, it doesn't follow that you have to know which mode of presentation is associated with any particular name; that a particular mode of presentation is necessarily associated with some name doesn't imply that you know that it is. To consider a more familiar example, even if what makes an animal a zebra is its DNA, it doesn't follow that someone who knows that it's a zebra will know anything about its DNA; metaphysical connections, however necessary, do not imply epistemological connections.<sup>35</sup>

With that said, we don't think our interpretation of the 'Sense and Reference' reply is too uncharitable to Frege. The philosophical community has only very recently come to have anything like a clear and distinct conception of the distinction between matters metaphysical and matters epistemological that makes the reply fail. So it isn't at all implausible that Frege (at least tacitly) thought that the close metaphysical connection he

35 Of course, in the case of a name, it might be impossible for *everyone* to be ignorant of the mode of presentation associated with it. For, if everyone were ignorant of the connection, it might be hard to see how that mode of presentation came to be associated with that name in the first place. But surely it's possible for *someone* to be ignorant of the connection.

posited between names and modes of presentation implied a close epistemological connection between them as well.<sup>36</sup>

We think we've demonstrated that Frege's 'Sense and Reference' reply to the objections to the name view relies on the claim that name types are individuated entirely by their semantic properties. But Frege's reply to the objections is incredibly brief and he doesn't say anything at all about how this claim about the individuation conditions of names is supposed to address the objections. So, although we think that we've offered a plausible reconstruction of Frege's reasoning in 'Sense and Reference,' it is possible that there is some other, equally good or better reconstruction. But, be that as it may, if there is such an interpretation, we don't think it will be one on which Frege is *rejecting* the name view in the opening paragraph of 'Sense and Reference.' For we *do think* that we've demonstrated that there's no genuine reason to think he does so and that the hypothesis that he rejects the name view depends upon an incredibly strained reading of the text. Moreover, the similarities between the relevant *Begriffsschrift* and 'Sense and Reference' passages make it very untenable that Frege is tacitly up to radically different things in these works; and Frege's continued assertion of the name view, as discussed in Section II, makes it completely untenable.

As stated above, Frege's 'Sense and Reference' reply to the objections to the name view of identity statements is incredibly brief. And this raises the question of why Frege goes to the trouble of raising the objections but doesn't bother to give a fuller reply. One answer to this question is surely that Frege takes himself to have already answered the objections in the *Begriffsschrift*; and, although the *Begriffsschrift* reply and the 'Sense and Reference' reply are somewhat different, they are similar enough that Frege may not have taken note of their differences. But this answer

36 In note 33 above, we discussed the possibility that Frege's 'Sense and Reference' reply involves the claim that sameness of mode of presentation type *as well as sameness of object type* is required for sameness of name type, rather than (as we interpret him in the text) the claim that sameness of mode of presentation type is necessary and sufficient for sameness of name type. In the note we gave reasons that favor the interpretation in the text, but we also said that in the end it doesn't matter to our discussion if the alternative interpretation turns out to be correct. If the alternative interpretation is correct, then obviously we will need to retract our claim that, for Frege, sameness of mode of presentation simply amounts to sameness of name. However, even on the alternative interpretation — and, indeed, on *any* reasonable interpretation — Frege is positing a very close metaphysical connection between which type a name token belongs to and the mode of presentation associated with it. Hence it's still plausible that this led him to think that *knowledge* of a name token's type (i.e., knowing which name the relevant sign is) implies *knowledge* of which mode of presentation is associated with it.

simply raises another question: why does Frege take the trouble to give a cursory reply to an objection that he has already answered?

In the final paragraph of 'Sense and Reference,' Frege appears to return to the questions about identity that begin the paper. He also uses identity statements when he first introduces the sense-reference distinction in 'Function and Concept.' And, as we will see, an examination of both these passages sheds light on why Frege's 'Sense and Reference' reply to the objections to the name view is so cursory.

## VII 'Function and Concept' and 'Sense and Reference'

Although Frege opens 'Sense and Reference' by raising concerns about the nature of identity, these concerns don't play any role in his ensuing discussion. After introducing the sense-reference distinction in the second paragraph, Frege spends a few pages discussing and clarifying these notions (for example, he says that in indirect discourse it is the sense, rather than the referent, that is under discussion; he says that senses are not to be confused with mental images; and he says that in imperfect languages some names have no referent). He then takes up what is the main topic of the paper: what are the senses and referents of declarative sentences? And it is only in the very last paragraph of 'Sense and Reference' that Frege makes any attempt to return the discussion to identity. As stated in Section I, although most philosophers know of the sense-reference distinction through 'Sense and Reference,' Frege first introduces the distinction in 'Function and Concept.' And his remarks on the distinction there are crucial to understanding both the main topic of 'Sense and Reference' and its closing paragraph, in which he appears to return to the issues about identity from the opening paragraph.

In the relevant passage from 'Function and Concept,' Frege first claims that a sentence's referent is its truth-value and hence that true sentences such as ' $2^2 = 4$ ' and ' $2 > 1$ ' have the same referent. He then raises a possible objection and offers a reply to it:

The objection here suggests itself that " $2^2 = 4$ " and " $2 > 1$ " nevertheless tell us quite different things, express quite different thoughts; but likewise " $2^4 = 4^{2^2}$ " and " $4 \bullet 4 = 4^{2^2}$ " express different thoughts; and yet we can replace " $2^4$ " by " $4 \bullet 4$ ," since both have the same referent. Consequently, " $2^4 = 4^{2^2}$ " and " $4 \bullet 4 = 4^{2^2}$ " likewise have the same referent. We see from this that from identity of referent there does not follow identity of the thought [expressed]. ('Function and Concept,' 13-14)

Frege goes on to give another example of two sentences that have the same referent but nonetheless express different thoughts: 'The evening star is a planet with a shorter period of revolution than the Earth' and

'The morning star is a planet with a shorter period of revolution than the earth.' The problem Frege is raising in 'Function and Concept' is *not* a problem about identity per se; although he uses ' $2^4 = 4^{2^2}$ ' and ' $4 \bullet 4 = 4^{2^2}$ ' to illustrate the problem, Frege also uses non-identity statements such as ' $2 > 1$ ' and the 'morning star' and 'evening star' sentences. Rather, the problem that Frege is concerned with in 'Function and Concept' is about sentences: how can two sentences that have same truth-value tell us different things or express different thoughts? As Frege indicates at the beginning of the passage, this problem arises for ' $2^2 = 4$ ' and ' $2 > 1$ ' simply because they have the same truth-value.<sup>37</sup>

Frege's solution to the worry about sentences in 'Function and Concept' is that it is the sense of a sentence rather than its referent that is relevant to the thought it expresses. He says:

" $2^4 = 4^{2^2}$ " and " $4 \bullet 4 = 4^{2^2}$ " have the same referent, but not the same sense (i.e., in this case: they do not contain the same thought). ('Function and Concept,' 15)

It is here that Frege first introduces the distinction between sense and reference and, importantly, he places a footnote at the end of the sentence:

I do not fail to see that this way of putting it may at first seem arbitrary and artificial, and that it would be desirable to establish my view by going further into the matter. Cf. my forthcoming essay "Sense and Reference." ('Function and Concept,' 15 n.F)

And, indeed, the main point of 'Sense and Reference' is to explain why his claims about the sense and reference of sentences are not 'arbitrary and artificial' and are in fact true. For, after dispensing with the preliminary discussion and clarification of the sense-reference distinction mentioned at the beginning of this section, Frege argues that customarily the referent of a declarative sentence must be its truth-value. However,

37. Many philosophers take the problems about identity raised at the beginning of 'Sense and Reference' to be at bottom problems about substitution failures: how can it be that co-referring names aren't always substitutable *salva veritate* or *salva significatione*? (See, for example, Salmon, *Frege's Puzzle*, 11-12 and 79-80.) It is true that, for Frege, the problem that he is raising in 'Function and Concept' is about the substitution of co-referring names; but that results from his idiosyncratic view that sentences are names that refer to truth-values, so for him two sentences that have the same truth-value just *are* co-referring names. And, insofar as Frege is concerned with the problem of substituting co-referring names that aren't sentences, it is because he is concerned with the problem about sentences and wants to show that the relevant sentences have the same truth-value. (See 'Function and Concept,' 13-14 and 'Sense and Reference,' 50, which are discussed in the text.)

given Frege's view that the referent of a complex expression is determined by the referents of its parts, this implies that sentences with the same truth-value should be substitutable *salva veritate*. But clearly this is not the case, and Frege spends the rest of the paper explaining why the cases in which substitution of sentences with the same truth-value fails constitute *special* cases and, hence, don't contradict his claim about what the customary referent of a sentence is. As stated above, Frege does attempt to clarify the notions of sense and reference in the first few pages of 'Sense and Reference'; but the main point of the work is clearly to defend his claims from 'Function and Concept' that the referent of a sentence is its truth-value while its sense is the thought it expresses.<sup>38</sup> And, importantly, none of the discussion after the opening paragraph in any way concerns questions about identity. However, Frege does attempt to return the discussion to identity at the very end of the paper, where he says:

Let us return to our starting-point.

If we found " $a = a$ " and " $a = b$ " to have different cognitive values, the explanation is that for the cognitive value,<sup>39</sup> the sense of the sentence, viz., the thought expressed by it, is no less relevant than its referent, i.e. its truth-value. If now  $a = b$ , then indeed the referent of " $b$ " is the same as that of " $a$ ", and hence the truth-value of " $a = b$ " is the same as that of " $a = a$ ". In spite of this, the sense of " $b$ " may differ from the sense of " $a$ ", and thereby the thought expressed by " $a = b$ " will differ from that expressed by " $a = a$ ". In that case the two sentences do not have the same cognitive value. If we understand by "judgment" the advance from the thought to its truth-value, as in the present paper, we can also say that the judgments are different. ('On Sense and Reference,' 50)

This closing passage is Frege's final word and a cursory glance at it makes it appear as if our interpretation of the opening paragraph may be in trouble. After all, when he returns to the question of why  $\lceil \alpha = \alpha \rceil$  and  $\lceil \alpha = \beta \rceil$  have different cognitive values, Frege says that, if  $\lceil \alpha = \beta \rceil$  is true, then the referent of  $\alpha$  is indeed the same as that of  $\beta$ , but claims that nonetheless the two statements can have different cognitive values. His claim as to why their cognitive values can differ makes no mention of the name view of identity statements, nor does he say anything about the need to individuate names by their modes of presentation. Hence

38 For a detailed discussion of Frege's view, in 'Sense and Reference' and elsewhere, that sentences refer to truth-values, see Tyler Burge, 'Frege on Truth,' in *Frege Synthesized: Essays on the Philosophical and Foundational Work of Gottlob Frege*, Leila Haaparanta and Jaakko Hintikka, eds. (Dordrecht: Reidel 1986), 97-154, at 97-123.

39 We translate '*für der Erkenntniswert*' as 'for the cognitive value' rather than 'for the purpose of acquiring knowledge.'

one enamored with the standard interpretation might be tempted to try to construe the final paragraph as a defense of the object view. Now we think the discussion in Section V shows that this is a completely untenable interpretation of the opening paragraph. But, more importantly for the discussion at hand, while a careful reading of the final passage does not indicate that Frege is in fact defending the name view, it does remove any temptation to think he's defending the object view. For, despite Frege's claim that he is returning to his starting point, and despite the fact that they are related, *the problem about cognitive value he considers in the closing paragraph and the problem about identity discussed in the opening paragraph are not the same problem.*

It is true that in the closing paragraph Frege says that, if  $\lceil \alpha = \beta \rceil$  is true, then the two names have the same referent. But he concludes from this that '*the truth-value of " $a = b$ " is the same as that of " $a = a$ "*' and says that in spite of *their truth-values being the same* 'the sense of " $b$ " may differ from the sense of " $a$ ", and thereby the thought expressed by " $a = b$ " will differ from that expressed by " $a = a$ ".' Notice that Frege in no way explains how the possible difference in sense between  $\alpha$  and  $\beta$  allows the two identity statements to express different thoughts and, hence, that what he says doesn't decide between the object view and the name view — that is, it doesn't determine whether he's using modes of presentation to defend the name view (as he clearly does in the *Begriffsschrift*) or whether he's instead using them to defend the object view (as the standard interpretation claims he does in the opening paragraph); and this is because his concern at the end of the paper is clearly to show that, despite the fact that the identity statements *have the same truth-value*, they can still express different thoughts. So, although Frege explicitly takes himself to be returning to the puzzles about identity with which the paper begins, and although he uses identity statements to illustrate a puzzle, the puzzle is not one of the puzzles about identity with which 'Sense and Reference' begins. Instead, it is the more general puzzle from 'Function and Concept' put in terms of cognitive value rather than thoughts: namely, how can two sentences with the same truth-value have different cognitive values?<sup>40</sup> And as is clear from 'Function and

40 The final paragraph of 'Sense and Reference' is not the only place in which Frege purports to discuss a worry particularly about identity but actually winds up discussing the more general worry from 'Function and Concept.' As mentioned in note 27, in the second volume of *The Basic Laws*, immediately after agreeing with Dedekind that the name view of identity statements is correct, Frege cites Thomae's objection to the object view: namely, that it implies that  $\lceil \alpha = \beta \rceil$  expresses the same triviality as  $\lceil \alpha = \alpha \rceil$ . However, although Frege has just asserted the name view of identity statements, his reply is not (as one would expect) that the name view avoids

Concept,' this puzzle does not have anything in particular to do with identity; it arises for ' $2^2 = 4$ ' and ' $2 > 1$ ' merely because they are both true, even though the latter doesn't contain the identity sign.

At the end of the last section, we raised the question of why Frege raises problems about identity at the beginning of 'Sense and Reference' but only gives an extremely cursory reply to these problems. In this section, we've seen that there is another question about Frege's discussion in 'Sense and Reference': when at the very end of 'Sense and Reference' he turns his attention to issues that began the paper, why does he raise the more general worries about sentences from 'Function and Concept' rather than the particular worries about identity from the beginning of 'Sense and Reference'? The answer to both these questions surely lies in what Frege's real concern in 'Sense and Reference' is. His main goal is not to give the proper account of identity; rather, it is to defend the claim from 'Function and Concept' that the referent of a declarative sentence is its truth-value. A crucial part of his defense is to concede that some of the apparent counterexamples really are counterexamples, but that in such cases what is customarily the sense of the sentence becomes its referent. And to do this, he needs to first get the sense-reference distinction on the table. But notice that he can't get it on the table in the same way he does in 'Function and Concept,' by drawing the distinction between the sense and referent of a sentence; *for it is precisely these claims about the senses and referents of sentences that he takes to be in need of defense*. So, in order to introduce the distinction, he brings up the puzzles about identity; but he introduces the distinction to solve the puzzle about sentences — not the puzzles about identity — by providing an indirect referent for sentences to refer to in belief ascriptions and other non-customary contexts.<sup>41</sup>

the objection that Thomae has raised against the object view so long as names are individuated by their modes of presentation. He does say that the explanation of why Thomae is mistaken is that the names can have different senses, but he goes on to say 'and that it is precisely the sense of the sentence — besides its referent, its truth-value — that determines its cognitive value' (§138, 140). Hence Frege clearly uses the fact that the names  $\alpha$  and  $\beta$  can have different senses to conclude that the sentences ' $\alpha = \alpha$ ' and ' $\alpha = \beta$ ' can have different senses, despite the fact that their truth-values are the same. So, just as in the final paragraph of 'Sense and Reference,' Frege begins by discussing the problems peculiar to the identity sign, but winds up discussing an entirely more general problem: namely, the worry from 'Function and Concept' about how two sentences with the same truth-value can differ in cognitive value.

41 The fact that much of 'Sense and Reference' is concerned with cases in which an expression fails to have its customary referent might seem to present a problem for

Notice also that it would be very natural for Frege to choose the puzzles about identity as a means of bringing up the sense-reference distinction. In the *Begriffsschrift*, Frege does not yet have an explicit distinction between an expression's sense and referent. Indeed, as we have already pointed out, he says later that his talk of *the content of a sentence* in the *Begriffsschrift* is ambiguous between thoughts and truth-values. But, as we saw in Section III, he does formulate a distinction between the mode of determination of a name and its referent when he discusses identity in the *Begriffsschrift*. And when Frege came to explicitly understand the distinction between sense and reference and its importance to his project, he certainly must have noticed that the distinction is exactly the distinction he used to defend the name view in the *Begriffsschrift*. Hence one would almost expect him to use the *Begriffsschrift* discussion as an uncontroversial means of bringing up the distinction, as indeed he does.

In short, we think that the brevity of Frege's reply to the objections to the name view in 'Sense and Reference' and his return at the end of the paper to the puzzle about sentences from 'Function and Concept' rather than the puzzle about identity that opened 'Sense and Reference' are best explained by the fact that Frege wasn't at all concerned about identity while writing 'Sense and Reference' and that, instead, he was entirely concerned with addressing the worries about sentences from 'Function and Concept.' Given the evidence adduced in the first six sections of this paper, it's clear that Frege was satisfied with his account of identity from

our interpretation of the text. After all, as we pointed out in note 22, in the *Begriffsschrift* Frege says that on the name view a name that flanks the identity sign refers to itself and, hence, *doesn't* have its customary referent. But if Frege is concerned in 'Sense and Reference' with cases in which an expression fails to have its customary referent, and if he still holds the name view there, then why doesn't he discuss the reference of names in identity statements? The answer is that, although Frege does spend much of 'Sense and Reference' discussing cases in which an expression fails to have its customary referent, *such cases are not his main concern*. As stated in the text, his concern is to defend his claim that the referent of a sentence is its truth-value; and he discusses cases in which he thinks an expression fails to have its customary referent *only* because they constitute potential counterexamples to his claim that the referent of a sentence is its truth-value. In each case, Frege claims that the relevant expression fails to have its customary referent, because this allows him rebut the potential counterexample. That is to say, his claims in 'Sense and Reference' about expressions that fail to have their customary referents are made in order to defend his claim about the reference of a sentence. The alleged fact that a name that flanks the identity sign doesn't have its customary referent is a consequence of the name view; but it doesn't serve any purpose in rebutting counterexamples to his claims about the referents of sentences, nor does it present any potential counterexamples to these claims.

the *Begriffsschrift* when he wrote 'Sense and Reference.' And, in the opening paragraph of 'Sense and Reference,' identity is merely a device to introduce the distinction between sense and reference; but the device, once used, was discarded.

### VIII Conclusion: The Importance of Sense and Reference

We've presented lots of evidence that Frege doesn't take the objections to the name view to be fatal and, contrary to the standard interpretation, never abandons the view. Frege replied to the objections in the *Begriffsschrift* and there is every indication that his reply in 'Sense and Reference' is not intended to differ from his reply there. Indeed, in the preceding section, we've argued that the discussion of identity that opens 'Sense and Reference' is merely a means of introducing the sense-reference distinction so that Frege can get to the main point of the paper: defending the claim that the referent of a sentence is its truth-value. But this raises the question of why defending this claim is so important to Frege. However, if one looks at the role that 'Function and Concept' and 'Sense and Reference' play in Frege's developing thought, this question isn't so hard to answer.<sup>42</sup>

In *The Foundations of Arithmetic*, numbers are defined as the extensions of concepts; and this definition plays an important role in Frege's attempt to show that every arithmetic truth follows from the laws of logic. But if numbers are extensions of concepts, then what are concepts? Frege attempts to answer this question in 'Function and Concept,' which he explicitly says is concerned with supplementing the logic of the *Begriffsschrift*.<sup>43</sup> Frege's answer is that concepts are functions. But for the answer to be informative, we need to know what the arguments and values of the relevant functions are. Frege informs us that the arguments are objects and the values are truth-values — the True when the object falls under the concept, and the False when it doesn't. But Frege also recognizes that this claim, together with his claims that the referent of a

predicate is a concept and that the referents of the parts of a sentence determine the referent of the sentence itself, naturally leads to the idea that the referent of a sentence is a truth-value. In 'Function and Concept,' Frege introduces the notion of sense to explain how sentences with the same truth-value can 'tell us quite different things' and, in 'Sense and Reference,' he considers counterexamples to the claim that a sentence's referent is its truth-value. So Frege's defense of this claim in 'Sense and Reference' is crucial to his over-all project of grounding arithmetic in logic; for, if it turns out that the customary referent of a sentence isn't its truth-value, his explanation of what concepts are is incorrect and, hence, his account of what numbers are is incomplete.

We've argued that the standard interpretation of 'Sense and Reference' is completely untenable; but in Section VI, we indicated some of the reasons that it might have seemed compelling. However, we think that there's another reason that Frege's discussion of identity in 'Sense and Reference' has been so misunderstood. Despite the fact that Frege never abandoned it, the *Begriffsschrift* conception of the identity sign is wrong. To state that Hesperus is Phosphorus, for example, is to make a claim about planets, not names; and, although Frege thinks that introducing senses can save the name view, as we've seen, neither his *Begriffsschrift* reply nor his 'Sense and Reference' reply actually works. And, quite independently of the effectiveness of Frege's replies, it is still grossly implausible that identity statements are all metalinguistic. However, the reply to the worries about identity that the standard interpretation imputes to Frege is a good deal less implausible. Given that 'Hesperus is Phosphorus,' for example, expresses that Hesperus and Phosphorus are identical, there is a puzzle about how it could be both true and informative; and the answer that the names have senses and that the sentence asserts something about Hesperus and Phosphorus by expressing a thought that contains these senses is a good deal better than the answer that the sentence expresses something about the names. So, no doubt, another reason that the standard interpretation has been so appealing is that it imputes to Frege a better view than the one he actually held.

Moreover, once one comes to recognize the appeal of the view that identity statements express a relation between objects by expressing thoughts that contain modes of presentation of those objects, it becomes apparent that there are more general problems about substitutivity that are solved if a name's contribution to the thought expressed by a sentence containing it is its sense rather than its referent; and it also becomes apparent that the sense-reference distinction is important to related issues in the philosophy of mind about the individuation of thoughts. The sense-reference distinction has important ramifications for central questions in both philosophy of language and philosophy of mind; and,

42 Thanks to Mark Kalderon for directing our gaze to the forest.

43 In the preface, Frege says: 'It is my intention, in the near future, as I have indicated elsewhere, to explain how I express the fundamental definitions of arithmetic in my *Begriffsschrift*, and how I construct proofs from these solely by means of my symbols. For this purpose it will be useful to be able to refer to this lecture so as not to be drawn then into discussions which many might condemn as not directly relevant, but which others might welcome' (i).

on the standard interpretation, Frege is much closer to seeing these ramifications in 'Sense and Reference' than he is if he continues to hold the name view there. Indeed, if we are right that the distinction is introduced primarily to serve Frege's larger purposes in the philosophy of mathematics, Frege is less concerned with the issues about language and mind that the sense-reference distinction has come to so thoroughly permeate than the standard interpretation inclines one to think he is.

So we have to confess that, on our interpretation of the discussion in 'Sense and Reference,' Frege is much less aware of the wider importance of his distinction than he is on the standard interpretation. Be that as it may, we think that the evidence we've adduced leaves little doubt that the standard interpretation is wrong. This is not to deny the importance of Frege's distinction between sense and reference to our understanding of language and mind, although it is to deny that he was fully aware of its importance. In the end, however, we don't think that Frege winds up looking bad. The sense-reference distinction does constitute an important conceptual breakthrough and it is still *Frege* who is responsible for it. Indeed, the history of philosophy hasn't seen too many distinctions of such import; and to expect a single person to both formulate such a distinction and fully grasp its import is perhaps to expect too much.

*Received: March, 2000*

*Revised: October, 2000*