

Chapter 1

Introduction

1.1 The topic of the dissertation

This dissertation is about our ability to represent one another's thoughts. It is about that capacity, or set of capacities, in virtue of which I can represent you as knowing that the earth revolves around the sun; as believing that *The Brothers Karamazov* is more profound than *War and Peace*; as assuming (or at least hoping) that this dissertation will be worth reading; and so on. It is also, albeit only indirectly, about our ability to represent one another's reasoning. It is about that capacity, or set of capacities, in virtue of which I can represent you not just as (for example) believing something, but as believing it, and believing it reasonably, precisely because you believe something else. Since, in the fundamental case, such representations of your thoughts and reasoning are *judgments* about your thoughts and reasoning, the dissertation is, most fundamentally, about a certain subcapacity of the capacity to judge: namely, the capacity to judge that another thinker thinks some particular thing or reasons in some particular way.

Many philosophers will know my topic by a different name. To speak in one familiar philosophical idiom, then, this dissertation is about propositional attitude ascriptions, or attributions. I begin by writing of our ability to represent one another's thoughts and reasoning, though, because, on the view I will defend below, this ability is what ultimately underlies and grounds our use of propositional attitude ascriptions, i.e., of sentences like 'Galileo believes that the earth revolves around the sun'.

I will add that the dissertation is, and was conceived as, part of a larger project in the metaphysics of mind, and specifically in that part of the metaphysics of mind that was once called the theory of judgment and, more recently, the theory of thought. The larger project is meant to issue, ultimately, in a representation of the whole of the faculty of reason. Here, however, my aim is only to produce a representation of one of the central components of that faculty: our ability to represent one another's thoughts—which ability is intimately bound up with our ability to represent one another's reasoning.

1.2 Representing reasoning

My thinking on the topic of this dissertation begins with the following ideas, and is guided by them throughout.

Human beings are rational animals. Because we are rational, we are able to engage in self-conscious reasoning about the world. If I believe, for example, both that Socrates is a philosopher and that all philosophers are wise, I can reason from these premises to the novel conclusion that Socrates is wise. Equally, if I believe both that Socrates is a philosopher and that all philosophers are wise, and I acquire evidence that Socrates is not in fact wise, I can reason from this evidence to the novel conclusion that, contrary to my prior beliefs, either some philosophers are not wise or Socrates is not a philosopher.

We are also, however, and again because we are rational, able to engage in self-conscious reasoning about one another's self-conscious reasoning. If I believe, for example, both that Plato believes that Socrates is a philosopher and that he believes that all philosophers are wise, I can reason from these premises to the novel conclusion that he (probably) believes that Socrates is wise. Equally, if I believe both that Plato believes that Socrates is a philosopher and that he believes that all philosophers are wise, and I acquire evidence that he does not in fact believe that Socrates is wise, I can reason from this evidence to the novel conclusion that, contrary to my prior beliefs, (probably) either he does not believe that Socrates is a philosopher or he does not believe that all philosophers are wise.

How, we might ask, is our ability to engage in self-conscious reasoning about the world related to our ability to engage in self-conscious reasoning about one another's self-conscious reasoning?

A striking fact about the above examples—a fact that suggests that the question just posed is indeed a question worth asking—is that there is a close structural parallel between my reasoning about Socrates and my reasoning about Plato’s reasoning about Socrates. In the first case, I reason from the premises that Socrates is a philosopher and that all philosophers are wise to the conclusion that Socrates is wise. In the second, I reason from the premises that Plato *believes* that Socrates is a philosopher and that Plato *believes* that all philosophers are wise to the conclusion that Plato (probably) *believes* that Socrates is wise. So suppose we represent my reasoning, in each case, as an inference. Then we have both

- (1) Socrates is a philosopher
All philosophers are wise
Therefore, Socrates is wise

and

- (2) Plato believes that Socrates is a philosopher
Plato believes that all philosophers are wise
Therefore, Plato (probably) believes that Socrates is wise.

Here the structural parallel between my reasoning in each case is rendered visible in the fact that the sentences ‘Socrates is a philosopher’, ‘All philosophers are wise’, and ‘Socrates is wise’ appear both in our representation of my reasoning about Socrates and in our representation of my reasoning about Plato’s reasoning about Socrates; and they appear, moreover, in the very same order.

There is also, of course, an obvious and important difference between these two inferences, a difference already marked by my inclusion of the parenthetical ‘probably’ in the third line of the second. The difference is that the first inference is deductively valid, while the second is at best inductively strong. Still—and this is the crucial point—even if the second inference is one that is merely inductively strong, it is plausible that its strength is at least partly a function of the deductive validity of the inference schema

- (3) a is F
Everything F is G
Therefore, a is G ,

a schema that, it seems, is instantiated—though, apparently, in rather different ways (or, more literally, in different places)—in both of the above inferences.

There is, in fact, positive reason for thinking that schema (3) is indeed instantiated in inference (2). Thus, consider the following two ways of schematizing inference (2):

- (4) S believes that q
 S believes that r
Therefore, S (probably) believes that p
- (5) S believes that a is F
 S believes that everything F is G
Therefore, S (probably) believes that a is G .

Clearly, instances of the more abstract schema (4) are not guaranteed to be inductively strong. Here, for example, is a perfectly fine instance of schema (4) that would make for a terrible inference: Martin believes that grass is green; Martin believes that snow is white; therefore, Martin believes that dirt is grass. By contrast, every instance of the more determinate schema (5) is guaranteed to be just as strong as inference (2). So it is not just the relatively abstract schema (4), but also the more determinate schema (5), that characterizes the form of inference (2). More precisely: it is in virtue of bearing the more determinate schema (5), and not merely in virtue of bearing the more abstract schema (4), that inference (2) counts as being of a generally valid form.

One obvious consequence is that the form of the thought that Plato believes that Socrates is wise, for example, is given, not just by the relatively abstract schema

- (6) S ϕ s that p ,

but, rather, by the more determinate schema

- (7) S ϕ s that a is F .

In other words, to generalize: the form of the thought that S ϕ s that p is partially characterized by the form of the thought that p : the internal structure of the thought that p is also part of the internal structure of the thought that S ϕ s that p . But the only way to account for this fact—or so I contend—is to show that the thought that p is a component *part* of the thought that S ϕ s that p .

1.3 A theory of ascription

My central aim, to speak in something like the familiar philosophical idiom mentioned above (in §1.1), is thus to develop and defend an account of thoughts that ascribe propositional attitudes, thoughts of the form $\langle S \phi$ that $p \rangle^1$ —for example, the thought that Plato believes that Socrates is wise. For it is in thinking such thoughts that we represent one another's thoughts; and so it is in thinking such thoughts that we represent one another's reasoning. The account I will defend derives from—though, importantly, it is not identical with—the theory of indirect discourse famously proposed by Gottlob Frege (1892: 28/153–54). My account can be characterized in terms of two of its central theses. The first is that the truth-conditions of the thought that $S \phi$ that p depend on the thought that p . The second is the claim already introduced at the end of the preceding section: namely, that the thought that $S \phi$ that p has the thought that p as a component part. Since Frege himself accepts the first of these theses, but rejects the second, the result of combining them is a neo-Fregean theory of thoughts that ascribe propositional attitudes—a neo-Fregean theory of ascription, as I will call it.

I want to use this first, introductory, chapter to provide some general motivation for accepting each of these two theses. (I have, of course, just given some motivation for accepting the second; but I will give more below.) Thus, in §1.4, I will explain my motivation for accepting the first thesis, and so for preferring a broadly Fregean approach to the theory of ascription. My discussion there will draw on Gareth Evans's (1982: Chapter 1) well-known interpretation of Frege's concept of a thought (*Gedanke*), according to which there is a close connection between the concept of a thought and the concepts of ordinary (so-called) propositional-attitude psychology, concepts like *judgment* and *belief* (to mention just two of the most central). As I will explain below, if the first, Fregean, thesis of the neo-Fregean theory of ascription is combined with Evans's account of the concept of a thought, the result is positively guaranteed to give a correct account of the truth-conditions of thoughts that ascribe propositional attitudes. So far as I can see, the same can be said for no other extant theory of propositional attitude ascriptions.

¹I will usually refer to thoughts and thought-forms using ordinary that-clauses, as I do later in this paragraph. But where, for grammatical reasons, that-clauses will not serve, I will instead place the relevant expressions between angle brackets, as I have done here.

In §1.5, then, I turn to a discussion of the second thesis. There, I explain my reasons for preferring a specifically neo-Fregean approach to the theory of ascription. My claim there is that only by accepting the second thesis, and so only by accepting the whole neo-Fregean theory of ascription, can we provide a philosophically satisfying explanation of the fact (as, I claim, it is) that your act of representing another thinker's thinking or reasoning involves an exercise of the very same capacity (or set of capacities) that you would exercise in thinking the relevant thought, or engaging in the relevant piece of reasoning, yourself. In short, representing another thinker's thinking or reasoning is not so much a case of thinking or reasoning *about* her thinking or reasoning as it is a case of thinking or reasoning, as it were, alongside her (or trying to). Unlike the discussion in §1.2, however, my discussion in §1.5 will focus directly on our ability to represent one another's thoughts. There, and for the remainder of the dissertation, our ability to represent one another's reasoning will remain in the background.

Finally, in the last two sections of the chapter, I will briefly explain what I take to be the central challenge to the neo-Fregean theory of ascription, and then provide a quick chapter by chapter overview of the argument of the dissertation as a whole.

1.4 Motivating Fregeanism

The first, Fregean, thesis of the neo-Fregean theory of ascription is, to repeat:

the first thesis: The truth-conditions of the thought that S ϕ s that p depend on the thought that p .

My reasons for accepting this thesis, and so for adopting a broadly Fregean theory of ascription, are grounded in the following familiar considerations.

On the face of it,

- (8) the thought that Jerry believes that Mark Twain wrote *The Adventures of Huckleberry Finn*

and

- (9) the thought that Jerry believes that Samuel Clemens wrote *The Adventures of Huckleberry Finn*

have different truth-conditions. For Jerry might not know, might not even believe, that the name 'Mark Twain' was the pen name of one Samuel Clemens. So he might believe that Mark Twain wrote *Huck Finn*, while denying (perhaps even on the basis of that very belief) that Samuel Clemens wrote *Huck Finn*. In such a case, apparently, thought (8) is true and thought (9) is false; for Jerry believes that Mark Twain wrote *Huck Finn*, but does not believe that Samuel Clemens wrote *Huck Finn*. But if there are possible circumstances under which two thoughts have different truth-values, it follows that those thoughts have different truth-conditions. So thought (8) and thought (9) have different truth-conditions.

The basic Fregean move is, first, to take this appearance at face value (that is, to assume that thoughts (8) and (9) really do have different truth-conditions), and then, second, to point out that the thought that $S \phi$ s that p and the thought that $S \phi$ s that q would seem to have the *same* truth-conditions precisely when the thought that p and the thought that q are identical, i.e., when the two are one and the same. In other words, the difference between the truth-conditions of thoughts (8) and (9) is traced to the fact that

- (10) the thought that Mark Twain wrote *The Adventures of Huckleberry Finn*

and

- (11) the thought that Samuel Clemens wrote *The Adventures of Huckleberry Finn*

are distinct thoughts. But to take this view is to take the view that the truth-conditions of the thought that $S \phi$ s that p depend precisely on the thought that p . It is thus to accept the first thesis.

As I said above, the real advantage of the first thesis, and so of the broadly Fregean theory of ascription, is that it is easy to show that it is guaranteed to have (what are, on the face of it) the right results. That is, it is easy to show that the theory will assign to thoughts like (8) and (9) exactly the truth-conditions they seem to have. The argument depends, however, on a certain view of the identity-conditions of thoughts. My defense of this view of thoughts—which is controversial—will need to wait until Chapter 2. Here, I will simply state the view and draw out some of

its consequences. My suggestion, to be clear, is that the consequences of the view, taken on their own, give us good reason to hope that it can be defended, since the resulting theory of ascription is, again, *guaranteed* to give the right results—something that no other theory can hope to do.

The view of thoughts I have in mind was introduced, most explicitly, by Gareth Evans. On Evans’s interpretation of Frege:

the single constraint Frege imposed upon his notion of a thought was that it should conform to what we might call “the Intuitive Criterion of Difference,” namely, that the thought associated with one sentence S as its sense must be different from the thought associated with another sentence S' as *its* sense, if it is possible for someone to understand both sentences at a given time while coherently taking different attitudes towards them, i.e. accepting (rejecting) one while rejecting (accepting), or being agnostic about, the other. (1982: 18–19, his emphasis)

In other words:

the Intuitive Criterion of Difference: If it is possible for someone to ϕ that p without ϕ ing that q , then the thought that p is distinct from the thought that q .²

If we accept this Criterion, we can use facts about the relations of compossibility between propositional attitudes to reach conclusions about the identity conditions of the associated thoughts. For example, as we have seen, it is possible for someone to believe that Mark Twain wrote *Huck Finn* without believing that Samuel Clemens wrote *Huck Finn*. If we individuate thoughts in accordance with the Intuitive Criterion of Difference, it follows that the thought that Mark Twain wrote *Huck Finn* is distinct from the thought that Samuel Clemens wrote *Huck Finn*.

²Formally: $\forall p \forall q (\exists \phi (\Diamond \exists S (S \phi \text{ that } p \wedge \neg (S \phi \text{ that } q))) \supset p \neq q)$. In other words, we can establish the distinctness of two thoughts by considering just a single attitude towards those thoughts. (I assume, additionally, that one can ϕ that p only if one grasps, i.e., “understands,” the thought that p .)

It may also be worth mentioning that, although Evans initially states the Criterion as a conditional, as I have done here, he almost immediately (and without comment) restates it as a biconditional. The full biconditional version of the Criterion is highly controversial, and so I have opted to work with the weaker conditional version here. But for a defense of the biconditional version (which, for the record, I accept), see Irad Kimhi, *Thinking and Being*.

What is important, at this juncture, is simply the shape of the theory of ascription that results when we individuate thoughts in this way. Thus, suppose, first, that thoughts themselves are individuated in terms of propositional attitudes, in accordance with the Intuitive Criterion of Difference. And suppose, second, that the truth-conditions of the thought that $S \phi$ s that p —a thought that *ascribes* a propositional attitude—will depend on the identity of the thought that p , in accordance with the first thesis of the neo-Fregean theory of ascription. Then the truth-conditions of the thought that $S \phi$ s that p will differ from the truth-conditions of the thought that $S \phi$ s that q just in case the thought that p is distinct from the thought that q . Putting these points together, it follows that, if the attitude of ϕ ing that p is distinct from the attitude of ϕ ing that q , then the truth-conditions of the thought that $S \phi$ s that p will differ from the truth-conditions of the thought that $S \phi$ s that q .

What Fregean theories of ascription reflect, in other words, is the natural and plausible assumption that the truth-conditions of our ascriptions of propositional attitudes differ in ways that reflect differences between propositional attitudes themselves. In other words, we distinguish, for example, the truth-conditions of the thought that Jerry believes that Mark Twain wrote *Huck Finn* from the truth-conditions of the thought that Jerry believes that Samuel Clemens wrote *Huck Finn* because we distinguish *the belief* that Mark Twain wrote *Huck Finn* from *the belief* that Samuel Clemens wrote *Huck Finn*. What the Fregean theory adds to our ordinary understanding of these matters is a precise theoretical vocabulary for characterizing these distinctions. For, with the notion of a thought in hand, we can say that what distinguishes the truth-conditions of the ascriptions is just what distinguishes the attitudes. What distinguishes the belief that Mark Twain wrote *Huck Finn* from the belief that Samuel Clemens wrote *Huck Finn*, for example, is that—as we can now say—the two beliefs involve (are attitudes towards) different thoughts: the thought that Mark Twain wrote *Huck Finn*, on the one hand, and the thought that Samuel Clemens wrote *Huck Finn*, on the other. Similarly, what distinguishes the truth-conditions of the thought that Jerry believes that Mark Twain wrote *Huck Finn* from the truth-conditions of the thought that Jerry believes that Samuel Clemens wrote *Huck Finn* is that each set of truth-conditions depends on a different thought: the first on the thought that Mark Twain wrote *Huck Finn*, the second on the thought that Samuel Clemens wrote *Huck Finn*. In short, our ascriptions of propositional attitudes will have

different truth-conditions precisely where there are different attitudes to be taken, that is, precisely where there are two distinct relevant ϕ ings—two distinct propositionally contentful mental states of the relevant kind (two distinct beliefs, for example)—in the offing.

It may help to drive the point home if I point out that there is an even more abstract structure here, which the Fregean theory of ascription fills out in a particular way, but which might, in principle, be filled out in some other way. For what makes the Fregean theory so attractive is precisely that other ways of filling out this structure are, in general, exceedingly unlikely to give the right results. In fact, I think, *no* other way of filling out this structure will *guarantee* the right results—as, we have seen, the Fregean theory does.

In order to avoid some irrelevant complications, I will describe this abstract structure in linguistic terms—keeping in mind, however, that this way of putting the following points is not (for my purposes) the fundamental one. In these terms, we can say that the function of the that-clause ‘that p ’ in the ascription ‘ $S \phi$ s that p ’ is to pick out some entity. And one goal, at least, of a theory of ascription is to identify a type of entity in terms of which we can go on to explain the truth-conditions of the whole ascription (I mean, as always, and in accordance with the above, the truth-conditions the ascription seems to have—that is, we’re still taking the appearances at face value here). In other words, we can abstract from the first thesis of the neo-Fregean theory of ascription to a sort of schema of that thesis, according to which the truth-conditions of a sentence of the form ‘ $S \phi$ s that p ’ depend on the entity picked out by the that-clause ‘that p ’, where it is left open, for the moment, just what kind of entity is in fact picked out by the that-clause. On the Fregean view, of course, the entity in question is a thought, with thoughts individuated in accordance with the Intuitive Criterion of Difference. But perhaps another kind of entity could do the trick. And if it can, we might be able to save the appearances without being Fregeans.

The challenge facing any such alternative view, however, is that the entities it employs must give the right results when it comes time to compute the truth-conditions of ascriptions in accordance with our new, non-Fregean, version of the first thesis. In particular, whatever kind of entity the that-clause picks out, entities of that kind must be at least as finely individuated as the truth-conditions of the ascriptions themselves. Thus, it would seem, entities of that kind must be at least as finely individuated

as the propositional attitudes themselves. That is, wherever the sentences ‘ S ϕ s that p ’ and ‘ S ϕ s that q ’ have different truth-conditions (again, taking the appearances at face value), the entities picked out by their that-clauses must also be distinct. Call this *the Matching Condition*:³

the Matching Condition: If the two ascriptions ‘ S ϕ s that p ’ and ‘ S ϕ s that q ’ have different truth-conditions, then the that-clauses ‘that p ’ and ‘that q ’ pick out different entities.

(Again, we could formulate the Matching Condition so that it is about thoughts rather than sentences, but it’s a bit complicated, and the basic point is the same either way, so I’ll just stick with the linguistic formulation.)

It is well known, of course, that *propositions*—at least on the most dominant understandings of them, as, for example, sets of possible worlds, or even as structured entities built up out of objects and properties—are not up to the present task. That is, the view that the entities picked out by that-clauses are propositions leads to violations of the Matching Condition. The basic problem is that we seem to use propositional attitude ascriptions to draw distinctions that, according to the propositional view, we simply are not drawing at all. For, on the propositional view, the sentences ‘Jerry believes that Mark Twain wrote *Huck Finn*’ and ‘Jerry believes that Samuel Clemens wrote *Huck Finn*’ quite simply *say the same thing*, have the same truth-conditions, because their that-clauses both pick out the same proposition. Suffice it to say, then, that the adherent of the propositional theory of ascription cannot take the appearances here at face value.

Of course, the appearances *could* be misleading. So perhaps, ultimately, we will simply need to bite the bullet and say that, despite how things seem to us, the sentences ‘Jerry believes that Mark Twain wrote *Huck Finn*’ and ‘Jerry believes that Samuel Clemens wrote *Huck Finn*’ have the same truth-conditions. I think, however, that the above discussion reveals that matters are not quite so simple. The idea behind the Fregean approach, remember, is that our theory of ascription should guarantee that two ascriptions of propositional attitude will have different truth-conditions precisely where there are different propositional attitudes to be taken. Now, you might think that the propositionalist could just reject this idea. But

³I was helped to see the importance of this Condition, and its usefulness in motivating the Fregean theory of ascription, by Matthews 2007: 201–209.

rejecting it would require taking the view that there are distinctions between different propositional attitudes that we are simply not equipped to describe. For the view would have to be that, while the belief that Mark Twain wrote *Huck Finn* is one thing, and the belief that Samuel Clemens wrote *Huck Finn* is another, the ascriptions ‘Jerry believes that Mark Twain wrote *Huck Finn*’ and ‘Jerry believes that Samuel Clemens wrote *Huck Finn*’ say the same thing, have the same truth-conditions. So which belief do they ascribe to Jerry? The belief that Mark Twain wrote *Huck Finn*? Or the belief that Samuel Clemens wrote *Huck Finn*? Or some third, as yet unmentioned, belief?

A really stubborn propositionalist might retreat, at this point, to the view that our ordinary understanding of the propositional attitudes themselves is just as wrong as our ordinary understanding of propositional attitude ascriptions. But it would, I think, be much more natural, and more plausible, to conclude instead—and against the propositionalist—that, whatever a “propositional attitude,” as understood by the propositionalist, is, it is not a propositional attitude. That is, for example, no matter what attitude one takes towards the *proposition* that Mark Twain/Samuel Clemens wrote *Huck Finn*, the result is not a *belief*—either that Mark Twain wrote *Huck Finn*, or that Samuel Clemens wrote *Huck Finn*, or anything else. In short, *beliefs* (for example) are just more finely individuated than the propositionalist’s “propositional attitudes.”

To put it another way: You can say, if you like, that, whichever belief you have (that Mark Twain wrote *Huck Finn* or that Samuel Clemens wrote *Huck Finn*), you are thereby related to the same proposition, namely (I suppose) the proposition that Mark Twain/Samuel Clemens wrote *Huck Finn*. (In fact, I see no reason for the Fregean to reject this claim, so long as it can *also* be true that each belief is a relation to a different *thought*.) And you can say, if you like, that, in employing either of the *ascriptions* ‘Jerry believes that Mark Twain wrote *Huck Finn*’ and ‘Jerry believes that Samuel Clemens wrote *Huck Finn*’, you thereby claim that Jerry stands in the relevant relation to the same proposition. (I see no reason for the Fregean to reject this claim, either, so long as it can also be true that, in using each ascription, you thereby claim that Jerry stands in the relevant relation to a different *thought*.) Actually, if you want to be really stubborn, you can even say, if you like, that there is only one “belief” here, the “belief” that Mark Twain/Samuel Clemens wrote *Huck Finn*; that is, you can insist that Jerry has exactly the same total set of “beliefs” in each of the relevant

cases. And you can apply that way of speaking to the relevant ascriptions of those “beliefs” as well. Whatever you say, though—however you want to use the relevant terminology—what remains is that, across the relevant cases, there is *some* difference in Jerry’s mental state, and he will think and act accordingly differently depending on which mental state he is in. And the difference in question is one that we usually—outside philosophical theory, at least—describe as a difference in *belief* (or, as the case may be, as a difference in *knowledge*); and it is one that we usually characterize, in practice, by using different propositional attitude ascriptions, with different truth-conditions.

I myself will say—speaking, of course, in the Fregean way—that the difference in mental state, and so the difference in subsequent (actual or potential) thought and action, is traceable to a difference between distinct *beliefs*. Equally, I will say that it is a difference in *what* is believed, and I will say that what is believed is a *thought*. What I am trying to suggest here is that, while this way of speaking may not be universally accepted, the view I mean to be expressing by means of it is. In short, the idea that there is simply no difference in mental state here is so implausible that no philosopher has been willing to endorse it. (Recall that even Russell would insist that seemingly problematic cases involve “disguised descriptions;” that is, even Russell allows that there is a difference in mental state.)⁴

The real challenge facing the Fregean, it seems to me, is to provide a philosophically satisfying defense of the admittedly difficult and understandably controversial view that Fregean thoughts are individuated in terms of propositional attitudes, in accordance with the Intuitive Criterion of Difference. But, as is perhaps already clear, I think that, if this view of thoughts proves indefensible, the prospects for providing an adequate account of the truth-conditions of thoughts that ascribe propositional attitudes will be exceptionally bleak. As Evans himself goes on to say (1982: 19, his emphasis), and as I have, in effect, now argued at some length:

Frege *needed* this connection between his theoretical notion of sense

⁴Likewise, on Salmon’s (1986) neo-Russellian view, belief is construed as a three-place relation between thinkers, propositions, and “guises” under which propositions are presented to thinkers. Salmon thus agrees with the Fregean that the difference in mental state is a difference in belief; but he construes that difference, not as a difference in *what* is believed, but instead as a difference in *how* what is believed *is presented to* the believer. Again, though, the difference is merely terminological, and so even neo-Russellians like Salmon are Fregeans in the intended sense.

and ordinary propositional-attitude psychology if that theoretical notion was to help him solve the original puzzle about how sentences composed out of expressions with the same *Bedeutungen*⁵ [like ‘Mark Twain wrote *Huck Finn*’ and ‘Samuel Clemens wrote *Huck Finn*’] can have different cognitive values [i.e., express different thoughts].

In other words, the notion of cognitive value—and so the notion of a thought, the cognitive value of a sentence—is itself one that belongs to what Evans here calls “ordinary propositional-attitude psychology” (although it took a philosopher to clarify it and give it a name).

So I think we all have reason to hope that the Fregean account can be vindicated. In Chapter 2, then, I will defend the view that thoughts can be explained in terms of a particular propositional attitude, namely, *judgment*. More precisely, I will claim that a thought can be defined (where the result is what used to be called a real definition) in terms of the concept of a *judgment* and the (Aristotelian) concept of a *capacity*. In short, I will define a thought as a capacity to judge. (More precisely: a thought is a judgment in first potentiality. I will explain this more precise formulation in Chapter 2.) So the thought that Mark Twain wrote *Huck Finn* will be the capacity to judge that Mark Twain wrote *Huck Finn*, the thought that Samuel Clemens wrote *Huck Finn* will be the capacity to judge that Samuel Clemens wrote *Huck Finn*, and so on. One upshot of this definition is that the thought that *p* will be distinct from the thought that *q* just in case the judgment that *p* is distinct from the judgment that *q*. Admittedly, this conclusion does not quite suffice to establish the truth of the Intuitive Criterion of Difference. But it is plausible that, for any attitude ϕ , if the judgment that *p* is distinct from the judgment that *q*, then the attitude of ϕ ing that *p* will be distinct from the attitude of ϕ ing that *q*. Given only this additional assumption, the account provided below does suffice to provide a justification for the Intuitive Criterion of Difference.

For the purposes of providing an account of the truth-conditions of thoughts that ascribe propositional attitudes, this foray into the metaphysics of mind may seem like overkill. And perhaps it is. Indeed, if my aim here were *only* to provide an account of the truth-conditions of sentences that ascribe propositional attitudes, and of the thoughts they express—that is, if my aim were only to say how the truth-conditions of

⁵Evans has ‘Meaning’ for ‘*Bedeutung*’. The rather vexed question of how exactly to understand Frege’s notion of *Bedeutung* is the topic of Chapter 3, below.

such thoughts and sentences depend on the semantic features of their parts—I would probably simply stipulate that thoughts are to be understood as individuated according to the Intuitive Criterion of Difference, and leave the defense of the Criterion itself for another day. But that is not my only aim. It is also part—in fact, an even more important part—of my aim here to defend the *second* thesis of the neo-Fregean theory of ascription, that is, the claim that the thought that p is a component part of the thought that $S \phi$ s that p . Since this thesis, unlike the first, is explicitly a claim about the metaphysical nature of a particular kind of thought, we cannot assess it unless we know something about the metaphysical nature of thoughts in general. So the discussion of Chapter 2 serves not only to provide some additional elaboration and defense of the Fregean account of the truth-conditions of thoughts that ascribe propositional attitudes; it also serves to provide a foundation for the discussion of the second, specifically *neo-Fregean*, thesis of the neo-Fregean theory of ascription, which thesis it is the primary goal of this dissertation to defend.

But before we get too far ahead of ourselves: why think that the second thesis is worth defending in the first place? Again, I did provide some motivation for the thesis already in §1.2 above. But I think it will be worth saying a bit more, if only to help bring out the relevance, to the second thesis, of the account of thoughts proposed in Chapter 2.

1.5 Motivating neo-Fregeanism

The second thesis of the neo-Fregean theory of ascription, once again, is this:

the second thesis: The thought that $S \phi$ s that p has the thought that p as a component part.

My reasons for adopting this second thesis, and so for taking a specifically neo-Fregean approach to the theory of ascription, are much less familiar than the reasons I gave, in the last section, for adopting the first. The quickest way to bring them out is to note that a grasp of, say, the thought that Plato believes that Socrates is wise seems to require a grasp of the thought that Socrates is wise. In more familiar linguistic terms, the claim is that an understanding of the sentence ‘Plato believes that Socrates is

wise' requires an understanding of the sentence 'Socrates is wise'.⁶ But we can perhaps summarize the point this way: knowing what it would be for someone to believe that Socrates is wise requires knowing what it would be for Socrates to be wise.

But there are other, deeper, considerations in play here. As I explained in §1.2, my basic motivation for defending the neo-Fregean theory of ascription is my conviction that reasoning about another thinker's reasoning is an exercise of the very same capacity for reasoning that is exercised in reasoning, oneself, about the same subject-matter. Importantly, however, this conviction is not itself ungrounded. It has its source, rather, in a certain understanding of the nature of self-conscious judgment.

My ultimate reasons for adopting the neo-Fregean theory of ascription are thus as follows.⁷ First, a self-conscious thinker who judges that *p* also judges that she judges that *p*; and she does so, moreover, *in* judging that *p*. (So it is not that she first judges that *p*, and then, a moment later, also judges that she judges that *p*. Rather, while she does, in a sense, perform two judgments—the first-order judgment that *p* and the second-order judgment that she judges that *p*—she nonetheless performs only a single, unitary (though composite) act: she judges, self-consciously, that *p*. A self-conscious act is thus a unity of a first-order act with—at least in the best case—its subject's knowledge of it.)⁸ Second, the judgment that *I* judge that *p* and the judgment that *you* judge that *p*—or, more generally, the judgment that someone else judges that *p*—have the same form. In other words, what I predicate of myself in judging that *I* judge that *p* is *the very same thing* I predicate of you in judging that *you* judge that *p*: according to my judgments, each of us *judges that p*.

From these points it follows that the thought that *S* judges that *p* is internally related to the thought that *p*. For, by the second point, the judgment that *S* judges that *p* (an other-ascription) is internally related to the judgment that *I* judge that *p* (a self-ascription). By the first point, though, the judgment that *I* judge that *p* is internally related to the judgment that *p*. Finally, since the judgment that *p* is internally related to the *thought* that *p* (because the thought is what, in the act of judgment, is recognized

⁶For a nice elaboration of the point, and of its significance, see Kripke 2008: 257–58.

⁷In this paragraph and the next, I am indebted to the work of Sebastian Rödl (especially his 2007).

⁸I discuss these issues in more detail in my "Inferring as a Way of Knowing" (in progress).

to be true),⁹ it also follows that the judgment that S judges that p (the other-ascription) is internally related to the thought that p . And, since the judgment that S judges that p is internally related to the thought that S judges that p (for the same reason), it follows that the *thought* that S judges that p is internally related to the thought that p .

But, I argue, Frege's own theory of ascription denies this last conclusion: on his view, the thought that S judges that p is *not* internally related to the thought that p . If I am right, Frege is thus committed to the view that the thought that S judges that p can be explained without appeal to the nature of the thought that p . The reason is that, on his view, the thought that S judges that p is composed of a way of thinking of S , a way of thinking of the function $Judges(x,y)$, and a way of thinking of the thought that p . So the thought that p is related to the thought that S judges that p only externally, only in virtue of being the object of a way of thinking that is a component part of the thought that S judges that p . To take this view, however, is to take the view that the form of the thought that p does not characterize the form of the thought that S judges that p . Or so, again, I aim to show.

We will see in Chapter 5, however, that Frege's defenders disagree. That is, they think they can deny that the form of the thought that S judges that p is characterized by the form of the thought that p *without* denying that the thought that S judges that p is internally related to the thought that p . Their claim is that the way of thinking of the thought that p that is, on Frege's view, a component part of the thought that S judges that p is related to the thought that p in an especially intimate way. In their language, the ways of thinking of thoughts that serve as the senses of sentences embedded in the that-clauses of propositional attitude ascriptions (the *indirect senses* of those sentences) are *canonical* ways of thinking of those thoughts. What is special about canonical ways of thinking, on this view, is that there is, for them, a "road back" (as Frege's defenders put it) from reference to sense: a canonical way of thinking of an entity is such that it can be determined on the basis of the entity alone. In other words, if one knows which entity is in question, one knows what the canonical way of thinking of it is.

The mistake Frege's defenders make here is this: they assume that

⁹As Frege says, for example, in "My Basic Logical Insights" (reprinted in Beaney 1997: 323). I explain and defend this claim in detail in Chapter 2.

thought about thinking has the same form as thought about the non-thinking world. They assume, more precisely, that thought about thinking must be *about* thoughts in exactly the sense in which thought about the non-thinking world is *about* such things as tables and chairs. This is exactly the view I rejected when I said, in §1.3, that thinking or reasoning about another thinker's thinking or reasoning should be understood as a case of thinking or reasoning, as it were, *alongside* the thinker (or trying to). My basic point here is that you cannot represent another thinker's thoughts except by thinking, yourself, the thoughts you represent her as thinking; and you cannot represent another thinker's reasoning except by engaging, yourself, in the reasoning you represent her as engaging in. Of course, in representing another thinker as, say, judging that p (i.e., in judging that she judges that p), I do not myself *judge* that p . Nonetheless, I do exercise, in another way, the very same *capacity* I *would* exercise in judging that p . Similarly, in representing another thinker as, say, believing that p because she believes that q , I do not, myself, judge either that p or that q ; nor, therefore, do I judge that p *because* I judge that q . Nonetheless, I do exercise the very same inferential *capacity* I *would* exercise in judging that p because I judge that q ; I simply exercise it in a different way.

Admittedly, matters are somewhat more complicated than these objections allow. At least some of Frege's defenders want to save *something* of the view just expressed. Tyler Burge and Christopher Peacocke, in particular, insist that thinking the thought that $S \phi$ s that p involves thinking the thought that p . Their intention, in other words, is to accept the claims just set forth. The problem, then, to be more precise, is not that Frege's defenders *reject* these claims. It is, rather, that they have no way of *explaining* why they should be true. In particular, they fail to reveal the connection between the thought that $S \phi$ s that p and the thought that p as an *internal* connection, as a connection pertaining to the very *forms* of thought in question. They dimly see, I think, that there *is* such an internal connection. But they stop at endorsing it. They do not continue to the important philosophical work of *explaining* it, of providing an account that reveals the connection to be genuinely necessary. They fail, in other words, to identify the principle behind the truth—which, to their credit, they do recognize—that thinking the thought that $S \phi$ s that p involves thinking the thought that p .

It might seem, however, that an account of canonical ways of thinking would provide the needed explanation. For if the thought that $S \phi$ s that

p is internally related to the canonical way of thinking of the thought that p , and the canonical way of thinking of the thought that p is internally related to the thought that p , then the thought that $S \phi$ s that p is internally related to the thought that p . The explanation will fall apart, however, if we find that the concept of a canonical way of thinking, in terms of which the explanation is formulated, is itself bankrupt. And, as I will show in Chapter 5, it is. To be more precise, I find that the concept is empty, that it is a mere artifact of the extensionalism of Frege's formal system—or, to put it another way, of his assumption that any adequate *Begriffsschrift* must be fully substitutional. On the version of Frege's theory of judgment I develop here, however, this assumption is not only unfounded; it is positively mistaken. Thus, from the perspective of Frege's theory of judgment itself—as opposed to the perspective of the formal system he imposes on that theory—the notion of a canonical way of thinking is otiose. It is an attempt to bridge a chasm opened up by Frege's adoption of a formal framework that is ultimately unfit for elaborating his theory of judgment.

My fundamental motivation for undertaking the project of this dissertation, then, is to develop an account of thoughts that ascribe propositional attitudes that is consistent with the above-described view of the relation between self-conscious thinking and reasoning and self-conscious thinking and reasoning about self-conscious thinking and reasoning.

A key component of the account I will develop is, of course, the second central thesis of the neo-Fregean theory of ascription, the claim that the thought that $S \phi$ s that p has the thought that p as a component part. Importantly, however, interpreters of Frege have found reasons—grounded in his writings themselves—for thinking that this claim is radically un-Fregean, i.e., that it is straightforwardly incompatible with some of Frege's most fundamental philosophical commitments. Thus, in the final substantive section of this Introduction, I will introduce some specific obstacles to the neo-Fregean view—obstacles that, again, have their source in Frege's own writings—and explain, in the abstract, how I think they can be overcome.

1.6 A Fregean obstacle to the neo-Fregean view

One of the main obstacles confronting the neo-Fregean view is a Fregean objection to its second central thesis. The objection (stated in linguistic

terms, for the sake of familiarity) runs as follows. On Frege's view, to say that the truth-value of the sentence 'S ϕ s that p ' depends on the thought that p is to say that the thought that p is the *Bedeutung*¹⁰ of the embedded sentence ' p ' (if only when it is so embedded). If the thought that p were also the sense of ' p ', when ' p ' is appropriately embedded, then the thought that p would be *both* the sense *and* the *Bedeutung* of the embedded sentence ' p '. But that cannot be: the sense of an expression, on Frege's view, is a mode of presentation of its *Bedeutung*, and a mode of presentation cannot present itself. So we must say, as Frege does, that, when the sentence ' p ' occurs in the ascription 'S ϕ s that p ', it expresses its *indirect* sense, a mode of presentation or way of thinking of the thought that p . In short, the relation between Frege's concepts of sense and *Bedeutung* is incompatible with the conjunction of the two theses that are supposed to be central to the neo-Fregean view. The latter view is thus either false or—at best—misleadingly characterized: either it is false, because one of its two central theses is false; or it is not Fregean (not even "neo"-Fregean), because it requires a rejection of Frege's conception of the relation between sense and *Bedeutung*. Nor is the misleading characterization innocent: if the proposed view is incompatible with Frege's conception of the relation between sense and *Bedeutung*, a defense of it cannot rest on an appeal to Frege's account of these concepts. Instead, a completely new account of these concepts, and of their application in the philosophy of mind and language, is required.

The bulk of the dissertation (Chapters 3–6) will be spent developing an adequate response to this objection. The response begins by admitting that the objection is partly right. For I agree that, if we take Frege's concepts of sense and *Bedeutung* to be defined by the theses about them that he seems to accept, then the very definitions of those concepts do indeed rule out the possibility of the neo-Fregean view. I will show, however, that the neo-Fregean view is both plausible (indeed, more plausible than the classical Fregean view) and deserving of the label 'neo-Fregean'. It becomes plausible once we abandon Frege's concept of *Bedeutung* in favor of a pair of concepts—*reference* and *semantic value*, in the terminology I will employ—each of which does only a part of the work of the concept they replace. And it deserves the label 'neo-Fregean' because it still employs

¹⁰For reasons that will become clear momentarily, I leave Frege's term '*Bedeutung*' untranslated.

Frege's own concept of sense. That concept is available to the neo-Fregean because, as we will see, the conflict between the neo-Fregean theory of ascription and Frege's theses concerning sense and *Bedeutung* dissolves once we replace the concept of *Bedeutung* with appropriate concepts of reference and semantic value.

What makes this solution possible is that the conflict in the neo-Fregean view, as adduced above, is generated by the interaction of two of Frege's theses about *Bedeutung*. The first is that the truth-value of a sentence is determined by the *Bedeutungen* of its parts. The second is that the sense of an expression is a way of thinking or mode of presentation of its *Bedeutung*. By replacing the single concept of *Bedeutung* with the two concepts of reference and semantic value, we can, in a certain sense, isolate these two theses, and so keep them from interacting in the way that rules out the neo-Fregean view. Thus, on the neo-Fregean view, we replace the first thesis about *Bedeutung* with the thesis that the truth-value of a sentence is determined by the *semantic values* of its parts. (This is a version of the principle I call SUBSTITUTION.) And we replace the second with the thesis that the sense of an expression is a mode of presentation of its *reference*. (This is what I call Frege's Principle.) From this perspective, the classical Fregean view can be seen to involve an additional assumption, to the effect that the semantic value of an expression is always its reference. (I call this the Classical Assumption.) This assumption is buried in Frege's Janus-faced use of the term '*Bedeutung*', but, once it is revealed as an assumption, we can see that it is in need of justification. We can also see that the behavior of expressions in propositional attitude ascriptions (and other "opaque" contexts) arguably constitutes a counterexample to the assumption, and so a reason to give it up—or rather, more accurately, to apply it only to expressions in "transparent" contexts.

1.7 Summary of the dissertation

The dissertation is divided into two parts. Part I treats of the general theory of thought (as I call it), a theory to which the subsequent account of thoughts that ascribe propositional attitudes—the theory of ascription—is thus answerable. Part II then explains and defends the neo-Fregean theory of ascription.

The central task of Chapter 2, as I have already said, is to defend the

Intuitive Criterion of Difference, i.e., the claim that, if, for some attitude ϕ , it is possible for someone to ϕ that p without ϕ ing that q , then the thought that p is distinct from the thought that q . I will defend the Criterion by providing an account—in fact, a definition—of thoughts in terms of a particular propositional attitude, namely, judgment. Specifically, I will define a thought as a certain sort of capacity to judge: an Aristotelian first potentiality. Correlatively, I will define a belief as a related but different sort of capacity to judge: an Aristotelian second potentiality.

As I understand these ideas, a first potentiality is a capacity the actualization of which is itself a capacity to perform some particular act. So, for example, the capacity to speak French, understood as a first potentiality, is something possessed by every normal human being; it is, as we might put it, the capacity to *learn* (to speak) French. The capacity to *speak* French, understood as a second potentiality (or first actualization), on the other hand, is something possessed only by actual French speakers. It is a capacity acquired when one *learns* French. So, importantly, the capacity to *speak* French (a second potentiality) is itself the *actualization* of the capacity to *learn* (to speak) French (a first potentiality). And the capacity to speak French is itself actualized in actually *speaking* French. But, crucially, there is really just *one* capacity here, with two stages of actualization. The capacity is the capacity to speak French, which is actualized, first, in learning to speak French, and is actualized, second, and most fully, in actually speaking French.¹¹

In terms of this structure—the structure of a capacity with first and second actualizations—we can, given an appropriate act or activity, define two different capacities, the first and second capacities to perform the relevant act or activity. My account of thoughts, then, involves the application of this idea to the act of *judgment*. Thus, the thought that Socrates is wise, for example, is the (first) capacity to *judge* that Socrates is wise. To possess this capacity is, quite simply, to possess its component concepts (also capacities): the relevant way of thinking of Socrates (namely, as Socrates) and the concept of being wise. The *belief* that Socrates is wise is then the first actualization of this capacity, an actualization paradigmatically achieved through the acquisition of evidence to the effect that Socrates is wise. The actualization of the belief that Socrates is wise is then, finally, the judgment that Socrates is wise. So the belief is *exercised* or *actualized*

¹¹I am indebted here to Kosman 2013: 57–62.

in, for example, reasoning with it, where such reasoning consists, in part, in *judging* that Socrates is wise.

The upshot is that a thought is a capacity. In fact, a thought is a *complex* capacity: it is a capacity made of up component capacities, capacities that are exercised together in judging what the thought is a capacity to judge. In these terms, the central question of the dissertation is one about the component capacities of, for example, the thought that Plato believes that Socrates is wise. The central thesis (the second thesis) of the neo-Fregean theory of ascription is that the thought—the capacity to judge—that Socrates is wise is a component part of the thought—the capacity to judge—that Plato *believes* that Socrates is wise. Thus, the very same capacities that would be exercised in judging that Socrates is wise are also exercised in judging that Plato believes that Socrates is wise.

In Chapter 3, which makes up the second half of Part I of the dissertation, I introduce the philosophical and terminological framework in which the neo-Fregean theory of thought—and so the neo-Fregean theory of ascription—is to be developed. Here I motivate and explain the three-fold distinction between sense, reference, and semantic value and relate it to Frege's familiar two-fold distinction between sense and *Bedeutung*. I show, in particular, that the concept of semantic value can be explained without appeal to Frege's notion of sense. I also argue, drawing on the conclusions of Chapter 2, that the concepts of sense and reference, and Frege's famous distinction, can be explained without appeal to the notion of semantic value. With that work done, I claim, it is easy to see that Frege connects the concepts of sense and reference to the concept of semantic value in a particular way by employing the term '*Bedeutung*' to express both the concept of reference and the concept of semantic value. Since these concepts are distinct and independent of one another, however, it becomes apparent that the connection is questionable.

I turn in Part II to the theory of ascription. I begin in Chapter 4 by presenting, in some detail, both the classical Fregean and the neo-Fregean theories of ascription, as well as a third, "two-level" theory, employing the philosophical and terminological framework of Chapter 3 throughout. By presenting the three views in the same—and, I claim, a neutral—terminology, we are better able to compare their respective advantages and disadvantages. I close the chapter with a discussion of one particular advantage of the neo-Fregean over the classical Fregean and two-level views: the first, unlike the other two, I argue, is semantically innocent in

the sense made famous by Donald Davidson (1968: 108). That, I suggest, gives us a first reason to reject the connection Frege forges between the concepts of sense and reference and the concept of semantic value.

A well-known objection to the classical Fregean theory of ascription has it that any language correctly describable by the theory would have to be unlearnable.¹² Recently, however, a number of philosophers¹³ have suggested that we can answer this objection by introducing a rule that determines the indirect sense of an expression on the basis of its customary sense, its doubly indirect sense on the basis of its singly indirect sense, and so on, all the way up the infamous infinite hierarchy of senses. In Chapter 5, by way of providing some additional motivation for the neo-Fregean theory of ascription, I show that this attempt to climb Frege's hierarchy fails. As I explain there, our understanding of the proposed rule turns out to reside entirely in the very thing the rule was introduced to explain: our grasp of thoughts that ascribe propositional attitudes. I conclude that, in order to save Frege, we must save him from himself. We must reject a key, and so far unquestioned, feature of his formal framework: namely, his extensionalism, the view that the truth-value of a sentence is a function (only) of the references of its parts. It is this feature of his formal framework, I argue, that forges the important connection, mentioned above, between the concepts of sense and reference and the concept of semantic value.

In Chapter 6, I continue my defense of the neo-Fregean theory of ascription by responding to some influential objections, originally due to Tyler Burge (1979, 2004). In so responding, I sketch a novel conception of logical form, according to which the logical form of a sentence (or a thought) is—to put it roughly—*the way in which* its truth-conditions depend on the semantic features of its parts. I then apply that conception of logical form to thoughts and sentences that ascribe propositional attitudes, with the result that there are two distinct ways in which a capacity can occur in a thought. This idea, it proves, suffices to undermine the most important objection to the neo-Fregean view, which objection turns out to rest solely on the undefended assumption that the truth-conditions of a thought can never depend on the parts of the thought themselves, but must instead depend only on the objects and functions of which its parts are ways of

¹²The objection is Davidson's (1965).

¹³Burge (1979 and 2004), Kripke (2008), Peacocke (2009), and Parsons (2009).

thinking.

In the final chapter, Chapter 7, I argue that the neo-Fregean theory, as presented in Chapters 4 and 6, can be extended to provide an account of *de re* and *de se* ascriptions—ascriptions of the forms ‘*S* ϕ s of *a* that it is *F*’ and ‘*S* ϕ s that she herself is *F*’, respectively. The central problem facing the neo-Fregean theory of ascription, in this area, concerns the semantic role of the pronouns ‘it’ and ‘she herself’: in short, the neo-Fregean seems to be forced to choose between the view that the semantic values of the pronouns are the *references* of their antecedents and the view that their semantic values are the *senses* of their antecedents; but neither of these options is acceptable. I show, however, that the dilemma is false. We can escape it by refusing to assign any semantic values at all to the pronouns in *de re* and *de se* ascriptions. What seems to require such an assignment is the assumption that *de re* and *de se* ascriptions, like *de dicto* ascriptions (i.e., those of the form ‘*S* ϕ s that *a* is *F*’), ascribe complete thoughts, but thoughts of a special kind. I argue that we should reject this assumption. Instead, I claim, we must see *de re* and *de se* ascriptions as ascriptions of *predications* (ascriptions of *self*-predications, in the case of *de se* ascriptions), rather than ascriptions of complete thoughts. The result is that both *de re* and *de se* ascriptions are to be recognized as primitive forms of thought. I show that these forms of thought can be explained in conformity with the neo-Fregean theory of ascription developed in Chapters 2–4, and that the resulting theory of *de re* and *de se* ascriptions gives the right results.

Throughout, what the neo-Fregean view exploits is the idea that we represent the thoughts of others through our own representations of the world. Wherever the same representational feature shows up in different thoughts, the same singular term or predicate shows up in adequate regimentations or formalizations of the sentences that express those thoughts. The neo-Fregean view exploits this insight in order to earn us the right to the view that each of the following thoughts involves the concept (in the usual philosophical sense) of being *F*: the thought that *a* is *F*, the thought that *S* ϕ s that *a* is *F*, the thought that *S* ϕ s of *a* that it is *F*, and the thought that *S* ϕ s that she herself is *F*. It thus earns us the right to claim that the relations we see to obtain between these thoughts—relations reflected in our use of the letter ‘*F*’ throughout—are *logical* relations. In this way, it earns us the right to claim that what we have here are representations of the logical forms of these thoughts. It thus allows us to do justice to the

view of the relation between self-conscious reasoning and self-conscious reasoning about reasoning that was described, briefly, above.