Thought Without Language: an Interpretationist Approach to the Thinking Mind

DISSERTATION

Presented in Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy in the Graduate School of The Ohio State University

By

Michael Dean Jaworski

Graduate Program in Philosophy

The Ohio State University

2010

Dissertation Committee:

Neil Tennant, Advisor

William Taschek

Ben Caplan
Abstract

I defend an account of thought on which non-linguistic beings can be thinkers. This result is significant in that many philosophers have claimed that the ability to think depends on the ability to use language. These opponents of my view note that our everyday understanding of our own cognitive activities qua thought bestows upon those activities the propositional structure of sentences and the inferential norms of public linguistic practice. They hold that our attributions of thought to non-linguistic beings project non-existent structure onto the cognitive activities of those beings, and assess the beings’ activities according to standards to which the beings bear no responsibility. So, despite the complex neural and behavioral activities of many non-linguistic beings, my opponents hold that those beings are not properly described as thinkers.

To respond to my opponents successfully, one must not merely cite scientific and folk practices of thought attribution that permit thought to be attributed to some non-linguistic beings. My opponents’ insights might be taken to demonstrate a need to revise those practices, or to treat the attributions of thought to non-linguistic beings made within those practices as instrumentally valuable but technically false. Instead, my strategy is to acknowledge the language-like structure and norms of thought, and show that a non-linguistic being’s cognitive activities might nonetheless have that structure and be subject
to those norms. I identify seven features pertaining to the structure and normativity of thought---intensionality, extensionality, control, reflection, objectivity, conceptual composition, and the institution of standards---that those who share my opponents’ intuitions might deem necessary for any cognitive activity to count as thought. I argue that the motivations for the reflection requirement rest on confusions about the nature of entertaining a proposition, and I reject some interpretations of the institution requirement. But I accept the other five requirements, and a properly-interpreted institution requirement, and show how a non-linguistic being can meet them. On the account I defend, a non-linguistic being counts as a thinker given that it obtains information from at least two sensory modalities, behaves in ways that are not immediate, invariant responses to stimuli, has been through a learning process in which its cognitive system underwent modification to better conform to standards of truth and rationality and track flourishing-relevant categories, and displays a pattern of cognitive and behavioral activity in which a rational pattern can be found.

I am guided in this process by reflection on the practical purposes that are served by our folk practice of thought attribution, chief among them the coordination of rational social living. I argue that these purposes place strictures upon the nature of thought that could potentially conflict with the structure cognitive science might find in cognition, but that the folk notion requires no legitimization from science. Furthermore, which cluster of attributions correctly describes a subject is always partially determined by attributers’ varying conceptions of ideal rationality and interpretations of how to render odd behavior
rationally explicable. These commitments yield my approach to the theory of thought, “Non-Scientific Interpretationism”.

iv
Acknowledgments

I thank those current and former members of the Ohio State University Philosophy Department who have served as experts on my various committees over the years: Neil Tennant (advisor), William Taschek, Louise Antony, Kevin Scharp, and Ben Caplan.
Vita

1998……………………………………B.A. Philosophy, Bowling Green State University
1999-2007……………Graduate Teaching Associate/Lecturer, The Ohio State University
2007…………………………...Lecturer, Ohio University-Lancaster, Pickerington Center
2007-present……………………………Lecturer, The University of Texas-Pan American

Fields of Study

Major Field: Philosophy
Table of Contents

Abstract ......................................................................................................................................... ii

Acknowledgments ....................................................................................................................... v

Vita ............................................................................................................................................... vi

Chapter 1: Thought in Science, Philosophy, and Everyday Practice ........................................... 1

Introducing the Issue .................................................................................................................... 1

Thought in Everyday Practice ..................................................................................................... 8

The structure of interpretive practice .......................................................................................... 8

The structure and norms of the everyday conception of thought .............................................. 9

The justificatory standards of interpretive practice ................................................................. 12

The purposes of interpretive practice ........................................................................................ 18

Thought in Science ..................................................................................................................... 20

Behaviorism and Cognitivism ................................................................................................. 20

Classical Cognitivism and its Rivals ......................................................................................... 21

Scientific Practices and Thought Attributions ......................................................................... 24
<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Sections</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Thought in Philosophy</td>
<td>27</td>
</tr>
<tr>
<td></td>
<td>Conclusion</td>
<td>33</td>
</tr>
<tr>
<td></td>
<td>Chapter 2: A Defense of Non-scientific Interpretationism about the Thinking Mind</td>
<td>35</td>
</tr>
<tr>
<td></td>
<td>Introduction</td>
<td>35</td>
</tr>
<tr>
<td></td>
<td>First Choice Point: Underdetermination by Data</td>
<td>36</td>
</tr>
<tr>
<td></td>
<td>Second Choice Point: Underdetermination and Scientific Pragmatic Factors</td>
<td>42</td>
</tr>
<tr>
<td></td>
<td>Third Choice Point: Dogmatic Realism vs. Interpretationism</td>
<td>46</td>
</tr>
<tr>
<td></td>
<td>Pattern, Noise, and Interpretation as Discovery of Facts about Minds</td>
<td>54</td>
</tr>
<tr>
<td></td>
<td>Final Choice Point: Scientific or Non-scientific Interpretationism</td>
<td>57</td>
</tr>
<tr>
<td></td>
<td>Conclusion</td>
<td>67</td>
</tr>
<tr>
<td></td>
<td>Chapter 3: From Requirements for Thought to Arguments for Linguistic Priority</td>
<td>69</td>
</tr>
<tr>
<td></td>
<td>Introduction</td>
<td>69</td>
</tr>
<tr>
<td></td>
<td>Purported Requirements a Subject Must Meet to Count as a Thinker</td>
<td>72</td>
</tr>
<tr>
<td></td>
<td>From the Seven Requirements to a Requirement of Language Use</td>
<td>76</td>
</tr>
<tr>
<td></td>
<td>Considerations that Davidson Marshals in “Thought and Talk”</td>
<td>78</td>
</tr>
<tr>
<td></td>
<td>Considerations that Davidson Marshals in “Rational Animals”</td>
<td>95</td>
</tr>
<tr>
<td></td>
<td>Arguments for sub-conclusion 1</td>
<td>98</td>
</tr>
<tr>
<td></td>
<td>Arguments for sub-conclusion 2</td>
<td>107</td>
</tr>
<tr>
<td></td>
<td>Considerations from Davidson’s later essays</td>
<td>113</td>
</tr>
</tbody>
</table>
Chapter 1: Thought in Science, Philosophy, and Everyday Practice

Introducing the Issue

A squirrel darts across a yard, behind a tree, and then off through some bushes and out of view. Rover the Chihuahua races across the yard, stops at the base of the tree, and paws at it, barking and growling, her eyes fixed on the treetop. Rover thinks the squirrel is in the tree.

Or so we would say. And we would be justified in doing so, according to the norms of our everyday practice of thought attribution. We feel as comfortable attributing thoughts like this one to Rover as we do to adult humans.

Contemporary science agrees with our folk practice. Behaviorism fell out of favor in psychology, replaced by approaches that attribute cognitive states to explain behavior. The study of animal behavior has flourished due to this change. We have no other way to explain many behaviors of non-human subjects than by attributing thoughts, often even higher-order thoughts. For example, Hare et. al. describe a situation in which a chimp observes a room being baited with food, and observes whether or not another chimp has also witnessed the baiting.¹ Only dominant chimps will approach the food directly; subordinate chimps either will not approach the food at all (in cases in which the

¹ Hare et. al. (2000).
dominant chimp observed the baiting also) or will approach the food only when the
dominant chimp will not notice (in cases in which the dominant chimp was observed to
not witness the baiting). This seems to indicate not only thought, but higher-order
thought. A similar phenomenon was observed with rhesus monkeys by Flombaum and
Santos.² Cheney and Seyfarth note the ability of vervet monkeys to obtain and recall
knowledge about social organization, such as rank within the group’s hierarchy,
matrilineal relationships, and the identity of those with whom they have shared positive
interactions in the past.³ And Routley describes behaviors among several species of birds
that suggest deception and behavior modification in response to failures.⁴

Because contemporary science and everyday practice endorse attributions of
thought to languageless beings like Rover, one might feel comfortable claiming that
Rover has thoughts. But there are reasons for discomfort. One of these has to do with
the history of the everyday practice of thought attribution. Pre-philosophical humans did
not employ a clear distinction between causation due to impersonal forces and
intentional, agentive causation.⁵ Thoughts were attributed to explain most phenomena.
We now believe that most of the phenomena we once explained by attributing thoughts
should not be explained that way. Ancient people had no other or better way to explain
the changing of the seasons, the movements of celestial bodies, etc. than to find agency
throughout nature. That did not show those explanations to have merit. It is easy to
concoct explanations in terms of beliefs, desires, and intentions that will portray any

² Flombaum and Santos (2005).
³ Cheney and Seyfarth (1985).
⁵ I refer here to a version of animism that is stripped of the metaphysical baggage (e. g., the claim that soul
or spirit is in all things) that has been foisted upon it by interpreters from other cultures. See Wiredu (1990).
movement whatsoever as the result of rational processes. We know that this is a mistake in many cases, and we know that it is a mistake that humankind has been prone to make. This gives us reason to be cautious when considering whether we should take a given attribution seriously. We want a principled way to distinguish between those attributions that should be taken seriously and those that should not.

Another reason for caution has to do with the relationship between thought and language, and the fact that canine communication systems lack some key features of human languages. The nature of that relationship can motivate a worry that only attributions made to linguistic subjects have explanatory merit. So, attributions to Rover are inappropriate because Rover does not use a public language.

Our folk conception of thought portrays our cognitive processes as an internal analog of public language. The sentences of human languages endorse, challenge, and raise for consideration various propositions about the world, stand in logical relations to one another, are composed of meaningful components (words), and can be generated ad infinitum from a finite base of such components. So it is supposed to be with our cognitive processes, according to the everyday attributive practice that treats these processes as instances of thought. When we cognize, we take attitudes toward propositions, reason well or poorly from some thoughts to others, and combine concepts, potentially ad infinitum, into wholes of propositional form. Thus, conceived qua thought, our cognitive processes share both structural and normative features with language.

This parallel between thought and language makes sense, given that one important function of a language is to allow us to express our thoughts. It is natural to suppose that
language has the structure it does because it could not serve its purpose without mirroring the structure of our cognitive processes—a structure those processes possess independently of the fact that we are language users. But the epistemic situation with respect to the supposition has a striking feature: of the two phenomena, language and cognitive processing, it is the one whose structure is supposedly derivative (language) that we can actually be sure has the structure. The easy public accessibility of facts about linguistic structure,⁶ and the relative inaccessibility of similar facts about the structure of cognition, suggests that, to arrive at our folk understanding of our cognitive processing as thought, we have framed it with the structure of language. If this plausible (though probably not provable) story is correct, then thought is cognitive processing as it appears when viewed on the model of an internal analog of a language.⁷

So, we might see our folk concept THOUGHT as projecting structure onto cognition rather than capturing structure that is already there, and see terms like “belief” and “desire” as capturing socially constructed kinds rather than natural ones. In that case, one might wonder whether it is appropriate to treat a being without language as having cognitive processes with the structure of language. Perhaps projection of linguistic structure onto cognitive processes is only appropriate when we must treat our subjects’ cognitive processes as that which their utterances publicly express. At least, we might feel that projections onto the cognitive processes of non-linguistic subjects should not be

---

⁶ At least, basic facts about linguistic structure—such as the fact that sentences are composed of words—are easily accessible. I do not deny that some facts about linguistic structure are quite difficult to establish.⁷ Here, I have used “thought” to name a kind of cognitive processing with specific structural and normative features. I will use the term “thought” and its cognates this way throughout the dissertation. When I wish to refer to processing more generally, or without commitment to its structure, I will use “cognition” and its cognates.
taken as seriously as those onto the processes of linguistic subjects. The motivation for such an idea only increases when we recognize that we not only project structure onto cognition when we treat it qua thought, but also the norms of inference. But if these norms have their basis in linguistic practice, it might be inappropriate to hold the cognitive activities of non-linguistic beings accountable to them.

Canine communication systems lack some of the elements of human language that relate it to thought. The sounds made by dogs do not decompose into meaningful components that can be rearranged to generate new, meaningful wholes without limit.8 This limits the extent to which they can stand in logical relations to one another, as it prevents them from containing terms that express logical operations. Two calls could still capture logical relations to some extent. For example, Call One might be only appropriately used in circumstances incompatible with those in which Call Two is appropriately used. But the lack of logical sophistication in turn limits the extent to which canine sounds could be used to endorse, challenge, and raise for consideration various propositions about the world. Since canine communication systems lack these key thought-related features of human languages, dogs do not count as fully linguistic subjects.

So, philosophical reflection disrupts our comfort in making thought attributions to Rover in two ways. First, it reminds us that many thought attributions should not be taken as seriously as attributions to normal adult humans. Secondly, it questions the

---

8 Or, if they do, we have not yet found any reason to believe they do. I am using dogs as an example because they are a familiar intelligent, but perhaps non-thinking, non-linguistic animal. If they turn out to be linguistic after all, alternative examples are plentiful.
appropriateness of treating non-linguistic subjects’ cognitive processes as having the structure of language, and as subject to norms based in linguistic practice.

Confronted with such challenges, it is not enough to simply cite everyday practice or contemporary scientific practice, for our philosophical reflections might be taken to show that those practices stand in need of revision. We need an account that can distinguish, in a principled way, between thought attributions that deserve to be made in full seriousness and those that do not. And, if we want to attribute thoughts to Rover, as the flourishing of cognitivist accounts of animal behavior suggests we should, we need an account that does not invoke features of fully linguistic practice in setting a sufficient condition for serious thought attributions.

My goal in this dissertation is to defend an account of thought on which it can be attributed just as seriously to some non-linguistic subjects as it can to fully linguistic humans. So, I will need a principled distinction between thought attributions that can be taken as seriously as ones made to fully linguistic humans and those that cannot. And this distinction will need to be drawn so that the sufficiency condition for serious thought attribution can be met by some non-linguistic subjects.

I will accomplish this in three main steps. In Chapter 2, I will argue that we should adopt a theoretical approach to thought called “non-scientific interpretationism”. On an interpretationist approach, facts about what a subject thinks are partially constituted by relations between the subject and interpretive practice. This contrasts with realist and eliminativist approaches to thought; the former holds that facts about what a subject thinks are determined independently of the subject’s relations to interpretive
practice, the latter holds that the category represented by the term “thought” is empty. An interpretationist approach is non-scientific in case it treats everyday interpretive practice as capable of establishing the existence of thought regardless of whether cognitive science establishes thought as a legitimate natural kind. In Chapter 3, I will identify the primary reasons why a non-scientific interpretationist might restrict serious thought attribution to linguistic subjects. I will show that arguments based on these reasons are not decisive. Such arguments seek to identify aspects of thought that depend on a subject’s status as a linguistic practitioner. In Chapter 4, I will defend a sufficient condition for serious thought attribution that captures those aspects of thought without restricting serious thought attribution to linguistic subjects. So, I will establish non-scientific interpretationism as the best approach to thought, explain why non-scientific interpretationism is often taken to entail that thoughts can only be seriously attributed to linguistic subjects, argue that there is room to question the purported entailment, then demonstrate that there is no such entailment by offering a compelling account of mind on which it does not hold. In Chapter 5, I clarify some details of my account by contrasting it with the accounts of other interpretationists, and discuss the implications of my account for the scientific, philosophical, and everyday approaches to thought attribution.

My motivation for this project stems from the fractured nature of the contemporary literature on thought. I want to give an account that unifies the insights of contemporary cognitive science and philosophical tradition. More specifically, I want to show how one can accept the insights of non-scientific interpretationism while still meeting two other constraints. The first of these is agreement with contemporary science
that thought does not depend on language. The second is that the distinction between serious and non-serious thought attributions must be preserved, and preserved in a way that treats attributions to some non-linguistic subjects with equal esteem as attributions to linguistic subjects.

Thought in Everyday Practice

On the view I will defend, facts about what a subject thinks are partially constituted by facts about the subject’s relations to interpretive practice. Given that interpretive practice will be the backdrop for much of the discussion in this dissertation, I want to make plain how I conceive of that practice. Specifically, the basic structure of the practice, the conception of thought it employs (i.e., the structural and normative features of cognition as it is treated by the practice), the justificatory standards it places on attributions, and our purposes in engaging in the practice are all relevant to much of the discussion to come. I want to be clear that my goal in what follows is to describe the actual process by which we typically come to offer interpretations of others, and not to elaborate upon Donald Davidson’s conception of radical interpretation.\(^9\)

\textit{The structure of interpretive practice}

\footnote{Davidson (1973).}
Interpretation is a social practice in which the behavioral patterns of subjects are judged according to the standards of rationality. The most basic cases involve a being, the interpreter, judging the behavioral patterns of a being, the interpretee or subject. The interpreter attempts to fit the behaviors of the interpretee to a rational pattern; in so doing, the interpreter attributes various thoughts to the interpretee. If the interpretee is linguistic, the process will also involve assigning meanings to the interpretee’s utterances, which are understood to express some of the interpretee’s thoughts. The interpreter may then judge the extent to which the subject has acted rationally.

Because interpreters judge the overall rationality of behavioral patterns, and not just quality of inferences, standards of practical rationality are most directly relevant to this process. But standards of theoretical rationality are important to interpretation as well, since a subject’s ability to meet them contributes to its ability to meet standards of practical rationality. Thus, to give good interpretations of a subject, one must consider what goal states and epistemic states to attribute to the subject, and how successful the subject has been in forming appropriate subgoals and intentions to act. When the subject is linguistic, the interpreter may present the interpretation to the subject, inviting the subject to self-interpret for purposes of comparison.

The structure and norms of the everyday conception of thought

The nature of interpretive practice commits interpreters to treat subjects’ cognitive processes as bearing certain structural features, and as subject to evaluation in terms of
certain norms. Interpretation treats behaviors as subject to norms of rationality. So, many standards are in effect. Among the most important are: consistency of beliefs, transitivity of preference rankings, inferential soundness, attentiveness to one’s reasons, conformity of will to reason, and maximization of expected utility. Most interpreters would not recognize most of these labels, yet the norms they capture are implicit in everyday interpretive practice. We criticize subjects whose wills do not conform to reasons they acknowledge, we are frustrated by subjects who hold recognizably inconsistent beliefs, and we are befuddled by subjects whose preferences seem disordered.

Most of these norms are not applicable to just any conception of cognitive processing. For example, part of the definition of soundness for deductive arguments demands a proper relation between propositions that are premises and a propositional conclusion. So, for a cognitive process to count as inferentially sound, it must contain elements that have propositional form. Thus, because the norm of inferential soundness is in effect in interpretive practice, interpreters are bound to treat their subjects’ cognitive processes as involving manipulations of elements that have propositional form.

Furthermore, norms that demand consistency of beliefs, transitivity of preference rankings, and maximization of expected utility require that certain attitudes are taken to propositions. For those norms to be in effect, a subject must, for some propositions p, q, etc., believe that p, prefer that q, etc. Thus, because norms that demand consistency of belief, transitivity of preference rankings, and maximization of expected utility are in
effect in interpretive practice, interpreters are bound to treat the cognitive processes of
subjects as embodying attitudes of belief and preference to some propositions.

The norms of interpretive practice, then, commit interpreters to treating the
cognitive states of interpretees as propositional attitudes. For a practice to employ those
norms, it must see cognition as having that structure. To abandon the view of cognition
as involving propositional attitudes, then, is to abandon judgment of behavioral patterns
in terms of the standards of consistency, transitivity of preferences, maximization of
expected utility, and inferential soundness.

Beyond a requirement of propositional attitudes, norms of interpretive practice
seem to demand that interpreters treat interpretees as having further characteristics of
personhood. For example, a norm that demands attentiveness to reasons assumes some
cognitive mechanism that attends to some cognitive states and not others. Consider that
point alongside a principle of fairness for normative demands: to say that a subject S
ought to do x and not y implies that one takes S to have control over whether S does x or
y. Given that principle, the norm of attentiveness implies that interpreters treat
interpretees as having control over the direction of their attention. Similarly, the norm
demanding conformity of will to reason implies that interpretees have some form of
control over their intentions and actions. One might say that behaviors conform more or
less well to rational standards without implying any control on the subject’s part, but
there is little point to demanding that the will conforms to rational standards unless one is
going to treat the subject as capable of exercising some degree of control over its
behaviors.
For my purposes, the most important norms of rationality are those that demand consistency of beliefs, transitivity of preference rankings, inferential soundness, and maximization of expected utility. These are most important for my purposes because these are the norms that force interpreters to characterize the elements of interpretees’ cognitive processes as propositional attitudes, and to be a thought is to be an element of a cognitive process that can be (properly) treated as a propositional attitude. Because I aim to give a sufficient condition for thought, the norms that force us to treat cognitive processes as thought are of most interest to me. Norms that demand attentiveness to one’s reasons and conformity of will to reason are also important parts of our everyday interpretive practice, but they relate more directly to the treatment of a subject as a person than to the treatment of a subject’s cognitive process as a thought.

*The justificatory standards of interpretive practice*

In the last section, I discussed some norms that interpreters use to judge interpretees. But interpretation, like other social practices, is itself governed by norms. So, interpreters, as such, are accountable to another set of norms, norms that specify what counts as proper interpretive practice. Interpretations can be judged better or worse than others in various ways, and must meet certain minimal standards to be acceptable at all. Of course, to count as an interpreter at all, one must attempt to find rational patterns in the subject’s behavior, but there are more specific standards that govern how this process is to be conducted.
For one, an interpretation must take into account the subject’s epistemic position. For another, an interpretation should take into account facts about the subject’s long-standing preferences. Suppose I know that Fred is a football fan, and that his team lost today. I know that Fred is likely to know the outcome of the game, and that he would have preferred a different outcome. If I see Fred at the bar with a bottle of whiskey, I might interpret him as attempting to dull the pain the loss caused. But if I know that Fred has had no access to facts about today’s game, that interpretation is unacceptable. Furthermore, if I know that Daisy is aware of the game’s outcome, but does not have strong preferences concerning football, I should not explain her moping strictly in terms of her awareness of the game’s outcome. In short, an interpreter is expected to consider what a particular subject tends to prefer, and to what facts this subject has had access.

Additionally, a good interpretation will take into account which facts are likely to be most salient to the subject. Sometimes this will involve knowing the subject’s long-standing preferences, as in the case of Fred and his love for football. Generally, the salience of facts to a subject increases with the importance of those facts according to the subject, all else equal. Of at least equal importance is the spatiotemporal proximity of the subject to the states of affairs that instantiate various facts. A subject that fails to attend to the here and now will not be interpretable according to rational standards for long. Thus, good interpretations will typically render the subject as acting at least partially on the basis of facts about the subject’s current environment.

Furthermore, the interpreter must attribute a level of conceptual and logical sophistication appropriate to subject and situation. In this endeavor, simplicity is a virtue.
One should attribute as little logical complexity and conceptual subtlety as one can while still rendering the subject’s behavioral pattern highly rational. Again, the interpreter’s knowledge of the subject’s background will be relevant. Some subjects’ behaviors tend to be based primarily on facts about the immediate environment; other subjects frequently act only after attempting to draw detailed comparisons between their current situation and key events in their pasts. If I know I am dealing with a ruminant subject, a higher level of complexity will often be appropriate in my interpretations than when I deal with other subjects.

Another standard has to do with distinguishing attributions that deserve to be taken fully seriously from those that do not. For example, I often say that my computer is “Trying to open a warning message”, and that my fully linguistic friend is “Trying to get her work done”. In doing so, I am attributing an intention to my computer, and an intention to my friend. But the first attribution does not deserve to be taken with the same seriousness as the second. This is because we all know that a rational pattern can be found in the computer’s movements only because it has been designed to move as it does by other rational beings. So, when rational patterns are present in a subject’s movements only as artifacts of others’ rationality, everyday practice does not take attributions to that subject seriously. Also, everyday practice does not take seriously attributions to beings with limited nervous systems, or beings that seem to behave strictly on the basis of instinct. When a group of moths bash themselves repeatedly against a porch light, we might offer various rationalizations of their behavior—“They’re trying to fly toward the Sun”, for example—but we do not take these seriously. The moths are just
bashing into the light because they have an instinctive attraction to it, and they are not savvy enough to overcome their instincts. 10

The attribution of an intention to my fully linguistic friend is a paradigm case of an attribution that deserves to be taken with full seriousness. It is made of a fully linguistic subject. Everyday interpretive practice treats as a truism the claim that sentences express thoughts. Therefore, it cannot help but treat language users as thinkers.

An additional constraint on taking an attribution with full seriousness is that interpretations must attribute appropriate kinds of concepts to subjects. For example, “Rover thinks the squirrel is in the tree” is a better candidate to be taken with full seriousness than, “Rover thinks the squirrel transubstantiated into a chunk of wood”, or, “Rover thinks the squirrel is divorced”. These later interpretations are not taken seriously because the propositional content of the attitudes they attribute contains concepts we should not attribute to dogs. In most cases, it is inappropriate to attribute a concept when the subject has neither ability nor need to attend to the difference between cases in which the concept applies and cases in which it does not. For linguistic subjects, there can be exceptions to this rule based on the subject’s deference to others who attend to the relevant differences, or the fact that the subject is part of a linguistic community in which the relevant differences do receive attention. Common intuitions about cases like Hilary

---

10 An interesting kind of case involves attributions to groups of beings: “The Buckeyes are trying to stop the Gators from scoring”, for example. We often take these more seriously in everyday practice than we do attributions to computers, but perhaps not as seriously as attributions to individual fully linguistic humans. A good account of thought must apply in a plausible manner to these kinds of cases as well as ones that involve a linguistic or non-linguistic individual as subject. On my account, such groups are candidates for serious attributions.
Putnam’s Twin Earth\textsuperscript{11} seem to show that everyday practice is sensitive to facts about subjects’ environments and history of interaction with items that have various properties, as well. A good interpreter brings together these types of facts about a subject’s situation, and presumes a mass of standing beliefs likely to be held by the subject, adding or removing beliefs based on the subject’s behaviors.

For non-linguistic subjects, this is a matter of observing how the subject interacts with its environment. When the subject is linguistic, a good interpreter typically begins by attempting to place the subject within a linguistic community of which the interpreter is also a member. The interpreter proceeds by interpreting the utterances of the subject as obeying the semantic and pragmatic conventions of that linguistic community, and revises to accommodate idiosyncrasies as they arise.

This is a matter of comparing behavior with expectation and convention. It is typically \textit{not} a matter of radical interpretation. To some extent, we retain the basic abilities that we used to learn our first language, but we do not constantly continue to learn it as we interact with other linguistic beings. Rather, we assume expertise at some point, and only stop to question it when we are presented with evidence that suggests we follow a convention poorly. So, although the learning process is never completed, neither is it continuous.

Linguistic subjects provide particularly rich behavior to interpreters, in that they have the ability to offer interpretations of themselves---both their thoughts and their speech. Everyday interpretive practice privileges self-interpretations. A good interpreter

\textsuperscript{11} Putnam (1975).
will not quarrel with a self-interpretation without some overriding reason to do so. Though thoughts are, by their very nature, entities capable of expression in public languages, they are not always practically accessible. Though I do not endorse the image of first-person access that depicts it as involving a mind’s eye gazing over all the mind’s thoughts, seeing them as they are, this is no reason to deny the apparent fact that first-person access usually outruns third-person access. I say usually, not always, because at times others are able to point out discrepancies between our behavioral patterns and our self-interpretations that show that our self-interpretations ought to be revised. I am inclined to take the usual superiority of first-person to third-person access as a data point, rather than a point of contention; a lack of details as to how and why first-person access has this advantage is no reason to suspend judgment on the matter of whether it has an advantage. But regardless of someone’s stance on that issue, it remains a fact that everyday interpretive practice favors the first-person, and treats interpretations as flawed insofar as they disagree with self-interpretations without doing so on the basis of overriding facts.

A good interpreter must be sufficiently attentive to all possessed facts about the subject, and must bear in mind the extent to which relevant facts about the subject are possessed. In this respect, interpretation of a subject is no different than most attempts to understand a phenomenon. All else equal, interpretations based on more facts are better than those based on fewer. The ideal is interpretation with full knowledge of the background preferences, epistemic state, and tendencies of the subject to act. Interpretations can be judged both by how closely they match the ideal, and by how well
they make sense of the subject given the extent to which they diverge from the ideal. The former is more useful for judging the quality of the interpretation as an object, the latter for judging the quality of the interpretation as an act.

*The purposes of interpretive practice*

We have several purposes in interpreting others’ behaviors. We have explanatory reasons: we want to explain why subjects have behaved as they have, especially in cases when they have done unexpected, severely suboptimal, or anti-conventional things. We have epistemic reasons: we want to know what they hold true so that we can rely on them as authorities, and we want to be able to judge whether they are reliably treated as authorities. We have predictive reasons: we want to predict what subjects will do next. Each of these contributes to reasons having to do with regulation of behavior: we want to coordinate our own actions with others’ actions, whether cooperatively or competitively. Furthermore, when coordination is cooperative, we want to be able to set shared standards for reasoning and action to facilitate the cooperative process.

In addition to these reasons for engaging generally in interpretation of behavior, there are several reasons why we want to engage in an interpretive practice that has the specific features identified in the past three sections. We have a general descriptive reason: we want a vocabulary that allows us to talk about what is going on when we use language and interact with the world, despite our lack of easy access to facts about the neurological structures and processes that are causally responsible for such activities. We
have a more focused descriptive reason: we want to describe what is going on in any
given episode of language use or interaction with the world. These two reasons each
involve our having a way of describing our cognitive systems as they relate to the world,
which is facilitated by a system that assigns propositional content to neurological states
and processes.

Furthermore, the goal of coordinating behaviors is facilitated by understanding
others in terms of their capacities to interact with the public objects of a shared external
world. This end is well served by a taxonomy of internal states of individuals that
categorizes those states in terms of capacities to interact with objects in the world—that
is, by a taxonomy of internal states as attitudes toward propositional, externally-
individuated contents. When we are seeking to coordinate behaviors cooperatively, a
system of classification that individuates inner states and processes partially by their
relations to external factors provides a structure for the setting of shared standards for
reasoning and action that is not thwarted by the idiosyncrasies of each individual’s
internal cognitive organization, as it would be with a system based on internally
individuated states. It is crucial to our getting by with each other in the world that we
understand each other’s cognitive systems as they relate to the world. A major source of
value of the everyday practice of rationalization lies in the application of shared
standards of rationality to the behavior of others.

The fact that one of the goals of interpretive practice is cooperation bestows upon
the practice a long-term, prescriptive or injunctive purpose: we want to teach and
reinforce standards of rational cognition, action, and interaction. These include standards
for the tracking of commitments and entitlements, a key skill by which we fit one another’s actions and cognitive episodes within a normative scheme vital to our social interactions and practical decision-making. Thus, a goal of interpretive practice is to teach and reinforce the rules of the basic practice of rational social living. This basic practice can be supplemented by procedures that make its underlying structure and purposes more tangible: the practice of explicitly asking for and giving reasons for actions and commitments.  

---

Thought in Science

*Behaviorism and Cognitivism*

For most of the early 20th century, behaviorism was the dominant scientific approach to thought. Science could not study a ghost in a machine, but it could study the machine. And so it did. Behaviorism in science and behaviorism in philosophy of mind buoyed and supported each other until doubts emerged about both forms of behaviorism. More specifically, philosophical behaviorism proved incapable of providing adequate reductions of statements containing mental terms in purely non-mental language, and scientific behaviorism struggled to account for language learning in early childhood.

Furthermore, the rise of computer science and neuroscience provided alternative

---

12 For more details, see Sellars (especially 1954), Lewis (1979), Brandom (especially 1994 and 1995b), and Lance (1997).
13 See Chisholm (1957) and Geach (1957).
approaches to theorizing about thought that avoided behaviorism without collapsing into mere metaphysical speculation. Each of the new fields maintained the idea that we must study the machine, but neither focused entirely on behaviors. Computer science focused on internal functional organization, along with inputs and outputs. Neuroscience focused on the normally inaccessible workings of the nervous system.

With behaviorism beset by problems, and alternatives available that allowed “thought” to be credibly placed within scientific and philosophical parlance, cognitivist approaches to thought could begin. Philosophically, thoughts could be placed in one of two relations to neural states and processes, either “is identical to” or “is realized by”. Scientifically, internal functional organization or neural structure could be hypothesized to explain behavioral data, and experiments could be designed to attempt to determine which of the rival hypotheses was best.

**Classical Cognitivism and its Rivals**

The classical approach to cognition takes very seriously the idea that neural processing is structurally and functionally analogous to language. On the classical approach, to be in a cognitive state is to be properly related to a mental representation with a propositional content. For example, to believe that grass is green, some part of my brain must embody a mental state with the content “grass is green”, and I must stand in dispositions to process that state in ways characteristic of belief. Reasoning is supposed to be carried out within an innate Language of Thought, which operates according to its
own rules of syntax.15 Inferential moves that are captured linguistically using logical operator terms have parallels within the Language of Thought, and mental inferences are accomplished by passing information, encoded in the form of mental representations, through the appropriate operational process. So, any given act of reasoning is realized by a localizable neural process in which information is processed in accordance with rules of innate syntax for representational states. These rules are not supposed to be imposed on the neural processes by us as observers; rather, they are there to be discovered. Learning a first public language is a matter of learning how to map meanings of public terms onto mental representations. Facts about what a subject thinks are fixed by facts that determine the content of their mental representations, and facts about the rules of reasoning that the subject either has applied or is disposed to apply to those representations.

Rival cognitivist approaches are less traditional. They do not conceive of cognition as being so tightly analogous to linguistic inference. On connectionist or Parallel Distributed Processing models, no localizable neural event can be identified as the realizer of a mental inference.16 Mental representation is accomplished by means of associative networks. Because these must be vastly interconnected, mental inference is not a localizable process; instead, its occurrence is distributed throughout the associative network, and it consists of a redistributing of associative strengths among the various nodes of the network. To the extent that patterns reminiscent of language and formalized logical inference are present, this is only because the network of associative strengths has

---

15 See Fodor (1975).
16 See Clark (1993a).
come to approximate them, and not because the system is actually following rules of syntax or performing in accordance with inferential rules. The mechanism of cognition on this approach is purely associative. The only innate aspect of the system is the initial weighting of links between nodes.

Connectionist networks are not always treated as rivals to the classical model. The two can be treated as compatible, each describing the same brain activity at a different level of abstraction. On such approaches, the connectionist model is taken to describe lower-level realizers for the higher-level processes of the classical model. But insofar as connectionist models are treated as rivals to classical cognitivism, connectionists view cognition as less language-like than do the classical cognitivists. Although the nets’ activities can capture inferential patterns, they do not accomplish this by means of a mental analog of an operator symbol—a localizable structure that represents those patterns.

Still less language-like are views on which cognitive processes are taken to be accomplished by means of mental maps. On such approaches, mental representation is more pictorial than symbolic, more analog than digital. Representational states are taken to order the represented items in ways similar to the ordering of the aspects of the world they represent. Reasoning based on such representations will resemble navigation more than logical inference. Whether there can be thought in this broader sense without public language is not my concern.

---

17 See the debates among Clark, Fodor and Pylyshyn, Ramsey, Stich, and Garon, Smolenksy, and Stich and Warfield, in MacDonald (1995) for debate over the proper understanding of the relationships among Connectionism, Classical Cognitivism, and everyday attributive practice.

18 See Kitchin (1994).
Indeed, it seems obvious to me that there can be thought in some broader sense, i.e. cognitive activity in the most generic sense, without public language. Non-linguistic animals do get around in the world, do process information, and do solve problems. But I am interested in giving an account of *propositional attitudes* that allows them to be entertained by the non-linguistic, so the existence of other forms of non-linguistic cognition does not constitute victory for me.

*Scientific Practices and Thought Attributions*

The various approaches I have mentioned each contribute to scientific practices in which thought attributions are made. Because these practices are scientific, their justificatory standards and purposes are not entirely coextensive with those of the everyday practice of thought attribution. Instead, the standards and purposes of these attributions reflect their status as scientific posits. Within scientific practice, a descriptive vocabulary has to earn its place by allowing us to systematically and efficiently explain and predict behaviors. Normative concerns play no official role in the determination of whether a vocabulary has scientific merit. Surely, science does sometimes attribute normative statuses. A scientist might, for example, identify the role that a certain type of neural structure or process plays in successful cognitive episodes, and thus develop a notion of the proper function of the structure or process. Proper functions can in turn be used to set standards for how the various elements of a cognitive system ought to perform, but the connection between these standards and normative injunctions is
indirect, because we lack direct voluntary control over whether the standards are met. One can try to gain long-term influence over the extent to which one’s cognitive system functions properly, e.g. by studying good inferential patterns and habituating oneself to follow them, but there is little one can do in the moment to ward off episodes of functional impropriety. This is contrary to everyday practice, in which we assume of our fully linguistic peers that they have control over their actions, allowing for a direct link between descriptive and normative purposes of attributions. The upshot is that, for the descriptive vocabularies of scientific attributive practice, efficient prediction and systematic explanation are the goals; we need not worry about whether such a vocabulary could play a normative role, as well, as we must with the terminology of everyday attributive practice.

So, there is an everyday practice of thought attribution, and there are scientific practices in which thoughts are attributed. There are several possible relationships that might exist between these practices. It might turn out that the best scientific practice preserves or refines the everyday notion of thought, and establishes the everyday notion as one that picks out or closely tracks a natural kind, a kind that would be recognized by an ideal scientific practice. Alternatively, it might turn out that the best scientific practice does not even approximate the everyday notion of thought, and establishes the everyday practice as one that is not even on the track of a natural kind. Or it might turn out that the best scientific account of cognition will not include a thought-attributing practice at all.19

Furthermore, a philosopher who aims for an understanding of the notion of thought can choose one of two options when it comes to understanding the relationship between each of these practices and the target notion. One might hold that our best scientific take on a topic is to be privileged as our guide to the subject. On this approach, a philosopher treats “thought” as intended to refer to a natural kind. The various scientific attributive practices are treated as attempts to illuminate the nature of the items that fall under that kind. Most of the scientific practices are regarded as failed attempts, or ones likely to fail, at accounting for thought, so they are taken to lack authority about what thought is really like. Only the favored scientific approach is treated as having any authority over what thought is like. The everyday practice is regarded as a mere folk practice, perhaps a practically indispensable one, but one that might fall woefully short of capturing the truth about thought. The folk practice suffices to establish thought as a topic for scientists to investigate, but has no authority over what thought is like. Only if the best scientific account of thought preserves the folk conception of thought will the folk conception be legitimized as capturing the truth about thought.

On the other hand, one might note the vital social purposes of the everyday attributive practice, and argue that the indispensability of the everyday notion of a thought renders it credible, regardless of whether science finds that the notion describes a natural kind. The best scientific account might not be well-suited to serve all the social

---

20 This position is agreed upon by participants in the debate among Fodor, Lepore, and P. M. Churchland in McCauley (1996).
purposes that the everyday notion does; if not, we would still need the everyday notion in addition to the scientific notion. On this approach, the everyday notion of thought is treated as the primary notion of interest, and “thought” is treated as naming a social kind, a category that is tracked as part of an extra-scientific social practice. The social kind known as “thought” might or might not be coextensive with any natural kind. To the extent that a scientific practice gives an account of cognition that is potentially hostile to the achievement of the goals of the everyday practice, the scientific practice is not studying cognition *qua* thought, even if it utilizes the term “thought”.

Thought in Philosophy

With these distinctions between possible inter-practice relationships and different privileging of practices, we can place contemporary philosophical approaches to thought within four broad categories. Three of these treat “thought” as attempting to name a natural kind and succeeding or failing. In one category are views on which science is in the process of vindicating the everyday notion of thought by showing that it (at least roughly) corresponds to a natural kind, as in Classical Cognitivism. 22 In another category are views on which science is discrediting the everyday notion of thought by showing that the best scientific practice attributes a different notion of thought, so the everyday notion does not capture a natural kind, as in some accounts of Connectionism. 23

---

22 See especially Fodor (1975).
23 See Clark, Smolensky in MacDonald (1995).
third category are views on which science is discrediting all notions of thought, by
showing that cognition is best studied without a thought-attributing practice.\textsuperscript{24} 

In contrast to these three kinds of views are approaches that treat “thought” as
naming a social kind and not attempting to name a natural kind.\textsuperscript{25} Proponents of this
approach see the task of accounting for thought philosophically as independent from the
way in which scientific study of thought proceeds. The existence and nature of thought
are determined by everyday practice, not scientific practice.

Which of these views is most reasonably held? To answer this question, we must
remember what features everyday thought has and why we are interested in employing a
notion that treats cognition as having those features. And we must consider the extent to
which each of the four views on inter-practice relationships and privileging of practices
facilitates or threatens those interests. This method will allow us to assess whether
“thought” is more reasonably treated as attempting to refer to a natural kind or a social
kind without determining, in advance, which course scientific discovery is most likely to
take.

Recall that, according to the everyday notion, thoughts are propositional attitudes.
Thus, our everyday attributive practice treats cognition as having the structure of
propositional attitudes. Because we treat subjects’ cognitive processes this way, we are
able to assess them according to standards of rationality, including inferential soundness,
consistency of beliefs, transitivity of preference rankings, and maximization of expected
utility. The application of these standards as norms within a community serves the

\textsuperscript{24} See Churchland (1992), Stich (1983).
crucial purpose of bridging the gaps between our idiosyncratic inner workings, providing a shared set of rules for rational living. Recognizing each other as subject to commitments and entitlements facilitates coordination of behaviors and vastly increases the efficiency of our interactions with other community members and the world.

So, by treating one another as taking attitudes toward propositions, we fit one another’s actions and cognitive episodes within a normative scheme vital to our social interactions. Insofar as we revise away from our understanding of a thought as a propositional attitude, we threaten our ability to subject cognition to these standards, and so threaten to obstruct some purposes for which we employ a notion of thought in the first place. With this in mind, we can assess which position is most reasonably held: that “thought” attempts to name a social kind, or a natural kind.

Suppose that one treats “thought” as attempting to name a natural kind. Then one will see questions about the nature and existence of thought as empirical questions. What thought is like and which beings, if any, are capable of it will be matters for scientists to settle as they become more certain which structure an ideal science would find in cognitive processes. What does this mean for our everyday notion of thought? One of two things will happen: it will be reinforced and vindicated by classical cognitivism, or it will cease to play a role in scientific studies of brain and behavior.

We can assess treatment of “thought” as attempting to name a natural kind without knowing in advance whether classical cognitivism or one of its rivals will eventually hold sway in the scientific study of cognition. The mere fact that classical cognitivism could fail is significant. The standard for a descriptive vocabulary within
scientific practice is that it allows for more systematic, efficient explanation and prediction than its rivals, and classical cognitivism might not provide this. Because classical cognitivism could fail, the very move to treat “thought” as a natural kind term exposes us to the possibility that thought, understood in the everyday sense as referring to propositional attitudes, might turn out not to exist. In that case, we would have to regard propositional attitudes as convenient fictions employed to achieve everyday social goals.

Practically, this might not matter much---the everyday practice could continue as it always had. But one might regard this as telling against the idea that the everyday notion was attempting to name a natural kind in the first place. If the practice could continue to serve its social purposes unscathed despite the failure of one of its key terms to refer to a natural kind, one might well wonder why we ought to think that the term was attempting to refer to a natural kind. The notion of a propositional attitude seems well-designed to play a role in an everyday practice that serves a vital social purpose, a purpose that could not be played by notions that, in the case we are currently imagining, turned out to be more useful for scientific purposes. So, on the supposition that classical cognitivism will lose out to one of its rivals, the everyday notion of thought would not only fail to be employed by the best scientific account of cognition, but it would also be exposed as working at cross purposes to the best scientific account of cognition. I submit that the reasonable conclusion in this case would be that the everyday notion of thought was not attempting to name a natural kind, after all.

But what if classical cognitivism were to win out over its rivals? In that case, the everyday notion of thought would have proven to be valuable for science as well as for
everyday purposes. One might then conclude that the reason why the notion works so well for everyday purposes is that it carves nature at its joints, as we intended.

But things are not so simple. This way of thinking obscures the fact that the need for the everyday notion would have remained even if the everyday notion had failed to be useful within science. The everyday need for the notion of thought does not stand or fall with its scientific utility, so the fact that one notion might successfully fill both roles should not be taken to indicate that the everyday notion had aspirations to fill both roles.

By analogy, suppose the existence of immaterial substance had been discovered (somehow) shortly after the publication of *The Concept of Mind*. Such a discovery would not have shown that, contrary to Ryle’s arguments, mental discourse was about immaterial substances after all. It might have been tempting to think that it did, but that would have been a mistake. Ryle found uses for mental discourse that had nothing to do with a ghost in a machine; if there had turned out to be a ghost there, that would not have shown that mental discourse was intended to be about the ghost, after all. Mental concepts did not need to be validated by the existence of immaterial substance.

Similarly, since the everyday notion of thought has purposes that have nothing to do with science’s verdict on the structure of cognitive processes, the scientific verdict that those processes do indeed have the structure of propositional attitudes would not reveal the everyday term to have been an aspiring natural kind term, all along. The problem with the suggestion that classical cognitivism might validate “thought” as a successful natural kind term is that it assumes that “thought” stands in need of some validation that

---

26 Ryle (1949).
only identification with a scientifically-recognized kind can provide. But this ignores the fact that everyday interpretation is subject to normative pressures that scientific identification is not. The divergent pressures upon the two processes show that they are not aimed at the same purposes. The everyday notion does not depend for its legitimacy on whether or not it best serves scientific purposes, so classical cognitivism cannot validate it any more than a rival approach could invalidate it. The success or failure of classical cognitivism has no more to do with the existence of propositional attitudes than does the existence or failure to exist of immaterial substance.

One might acknowledge that the everyday notion of thought could be shown to have some practical legitimacy independently of the scientific study of cognition, yet worry that this does not secure the conclusion that thoughts actually exist. The short answer in response to this worry is that many kinds of things exist despite the fact that the kind is not a natural one. “Jade” does not name a natural kind---there are two scientifically distinct kinds of mineral that are properly called “jade”---but jade exists. Similarly, the existence of dinosaurs was threatened much more by the K-T asteroid impact than it ever was by the fact that the label “dinosaur” turned out to name a polyphyletic group of reptiles. I will grant that the worry deserves a longer response, one that does not rely so much on folk intuitions about what exists. To give the longer response requires more terminological background than I have provided so far, but it will emerge from the discussion in Chapter 2.27

27 See the section “Final Choice Point: Scientific or Non-scientific Interpretationism”.
For now, we can tentatively conclude that thought is a social kind, not a natural kind. If this tentative conclusion holds true, eliminativism is eliminated as a serious option in the philosophical debate about thought. Eliminativists eliminate thought by arguing that “thought” is an aspiring natural kind term that fails to refer to an actual natural kind. Thus, they assume that thought is supposed to be a natural kind. Furthermore, I doubt that one could view thought as a social kind and still view the category that “thought” refers to as currently empty, so I do not believe it is an accident that eliminativists have treated thought as an aspiring natural kind in their arguments. Of course, the conclusion that thought is not a natural kind also eliminates any realist or interpretationist views that depend on treating thought as a natural kind. So, a successful defense of the claim that “thought” names a social kind, and makes no attempt to name a natural kind, severely restricts the theoretical options available to one who wishes to give a philosophical account of thought.

Conclusion

In this chapter, I have established a goal for this dissertation: the defense of an account of thought on which thought can be attributed just as seriously to some non-linguistic beings as it can to linguistic humans. I have also laid out some background information to provide context for what is to come. Thoughts, in the everyday sense, are propositional attitudes. In the social practice known as interpretation, cognition is treated

as having the structure of propositional attitudes. It is part of our everyday practice to treat some non-linguistic beings as thinkers, but philosophical reflection has led to some deep worries about whether non-linguistic beings can think. The structure of propositional attitudes seems to be modeled after the structure of language, and we might be mistaken in believing that a non-linguistic being could have cognitive episodes with this quasi-linguistic structure.

My goal in the remaining chapters is to defend a view of the thinking mind on which some non-linguistic beings can be truly considered thinkers, just as linguistic beings can. In Chapter 2, I will argue that one type of view, non-scientific interpretationism, provides the correct approach to the thinking mind. My defense of non-scientific interpretationism expands upon the discussion in the “Thought in Philosophy” section of Chapter 1. By treating cognition as thought, we are able to subject our cognitive processes to standards of rationality, for the purpose of setting shared norms for social living. These norms allow us to better coordinate our actions with one another, and facilitate our ability to interact with the non-social aspects of our environments, as well. Largely because of this normative concern, thought is fundamentally a social kind, and its existence is established independently of the success or failure of scientific models of cognition that treat it as having the structure of thought. Furthermore, the variety of normative pressures allow for variability within the class of proper interpretations, a form of indeterminacy that can be resolved only given the contribution of someone who takes on the active role of interpreter.
Chapter 2: A Defense of Non-scientific Interpretationism about the Thinking Mind

Introduction

Within philosophy, there are three main kinds of approach to the thinking mind: realism, eliminativism, and interpretationism. Eliminativists hold that the category represented by the term “thought” is empty. Interpretationists hold that facts about what a subject thinks are partially constituted by facts about relations between the subject and interpretive practice. Realists hold that facts about what a subject thinks are determined completely independently of the subject’s relations to interpretive practice. My goal in Chapter 2 is to defend a version of interpretationism, non-scientific interpretationism, as the proper approach to the thinking mind.

I will accomplish this goal by presenting a series of choice points, eliminating some views from consideration as each choice is made until only non-scientific interpretationism remains. Not only will this process provide the reader with an argument that non-scientific interpretationism is the correct account of the thinking mind, it should also give the reader a fuller understanding of what I take a commitment to non-scientific interpretationism to entail, and, importantly, a more detailed account of why

---

30 See especially Davidson (1973) and Dennett (1971).
“thought” should be treated as intended to refer to a social kind rather than a natural kind. In Chapter 1, I gave a quick argument for the conclusion that “thought” names a social kind, not a natural kind\(^\text{32}\), and promised to discuss the matter in additional detail in Chapter 2. Securing the conclusion that “thought” names a social kind will eliminate eliminativism from contention to be the best theoretical account of the thinking mind, since defenses of eliminativism assume that “thought” aspires to name a natural kind but fails.

First Choice Point: Underdetermination by Data

My first step is to call attention to a basic commitment that distinguishes realists from interpretationists. Consider a situation in which an interpreter has amassed some data about a subject. The interpreter’s goal is to attribute various mental states to the subject, which involves assigning contents to the attributed states such that the subject’s activities are rendered rationally explicable. The interpreter will try to find a combination of attributions that allow us to see why a rational being would have acted as the subject has. Call any such combination of attributions an \textit{interpretation} of the subject.

According to a realist, there can be only one correct interpretation of the subject: the interpretation that attributes the thoughts that the subject actually has. The realist holds that there is a fact of the matter, determined independently of the interpretive process, as to which thoughts a subject has. Correctness of interpretation is a matter of

\(^{32}\) See the section “Thought in Philosophy.”
correspondence between the states that are attributed and the independently determined states of the subject.

Interpretationists, on the other hand, are free to acknowledge the possibility of multiple, equally correct interpretations. Because the interpretationist insists that thoughts are not fully determined independently of interpretive practice, correctness of interpretation cannot be treated as a matter of correspondence between attributions and independently determined thoughts. The relevant data about subjects limit the class of correct interpretations, but all things considered, correctness is a matter of working within the boundaries set by the relevant data in a way that achieves the goals of interpretive practice. These goals, specified in Chapter 1, are not merely descriptive, but prescriptive as well.

For a number of reasons, this opens up the possibility that the contribution of the interpretive process to the determination of thought will be variable rather than constant. Interpreters might reasonably disagree about which cluster of attributions best accomplishes the goals of interpretation, and such disagreement might be incapable of resolution by appeal to facts. This could be because the interpreters differ legitimately in their general assessment of what counts as best accomplishing the goals of interpretation. Or the interpreters might agree as to what counts as best accomplishing the goals of interpretation in general, but find themselves in circumstances such that each

---

33 See the section “The purposes of interpretive practice” for details.
34 This could occur because one interpretation renders the subject more rationally explicable to an interpreter who favors maximin strategies for choice under uncertainty, whereas another interpretation renders the subject more rationally explicable to an interpreter who prefers to take more risks.
must interpret differently than the other in order to follow the agreed-upon method for accomplishing the goals of interpretation. 35

Because interpretationists are capable of recognizing multiple interpretations as equally correct, and realists are not, one way to argue for interpretationism is to show that we have good reasons to believe that there can be multiple correct interpretations, and no reason to think that there ought to be only one correct interpretation in every case. My burden for this section and the two that follow will be to successfully pursue such a strategy. There are many features of the interpretive process that might lead a theorist of thought to conclude that there can be multiple correct interpretations of a subject, and as we encounter and acknowledge each of them, various realist views will be eliminated as acceptable accounts of the thinking mind.

With this in mind, consider again the interpreter, in possession of data about a subject, and poised to attribute thoughts that will render the subject’s activities rationally explicable. In any such scenario, there are obvious epistemic limitations with which the interpreter has to contend. No matter how skilled the interpreter is, she usually will not have access to all the relevant facts about the subject’s activities. Subjects typically act in ways that reflect not just their perceptions of immediate circumstances, but also lessons learned from events long past, events of which the interpreter might be completely unaware. And even if an interpreter had access to all the relevant facts, these facts might be too numerous for the interpreter to successfully keep them all in mind for the purpose

35 This would indicate a difference in the needs of the interpreters’ communities, and, because of that difference, a difference in the details concerning how best to accommodate the prescriptive goals of interpretation.
of holistic consideration of them. Given these epistemic limitations, a wide range of attributions will be compatible with the data the interpreter has about the subject, each of which corresponds to a different way of filling in the epistemic gaps.

This merely practical kind of indeterminacy of proper interpretations is of little use for the task of supporting interpretationism against realism. It reflects the fact that there is relevant information that the interpreter is unable to consider. This does nothing to counter the realist’s claim that, were the epistemic gaps filled in, there would be only one correct pattern of attributions to be made.

To ensure that my arguments for the possibility of multiple correct interpretations do not rest on this merely practical indeterminacy, we should consider a version of the interpretive scenario in which we have idealized away from these practical epistemic limitations. Imagine that the interpreter is somehow in possession of all the data about the subject that is relevant to interpreting the subject. These data will include non-relational facts about the subject, including ones not normally available to everyday interpreters, such as specific facts about the subject’s brain state. And these data will include any relational facts needed to assign a wide rather than narrow content. They will include facts about the subject’s relations to all of the following: history, status as a member in various linguistic/non-linguistic communities, and current environmental context. And they will include knowledge of the subject’s dispositions. Additionally, suppose the interpreter is capable of considering sets of attributions that offer complete descriptions of the subject’s overall cognitive state, so that the interpreter not only has all the relevant information, but will be able to deal with it in a truly holistic manner. In
such a scenario, we have closed the gap between the epistemic position of the interpreter
and one class of potential metaphysical determiners of the subject’s thoughts—the class
of facts about the subject that a realist could find acceptable as metaphysical determiners
of a subject’s thoughts.

Suppose the idealized interpreter attends to all these facts about the subject, and
seeks to give an interpretation of the subject. Suppose additionally that the interpreter
tries to base the interpretation only on full knowledge of the data, and with strictly
descriptive intentions. That is, the interpreter will appeal to neither the pragmatic
features that normally affect theory choice, such as elegance and parsimony, nor the
prescriptive or normative concerns of the everyday practice of interpretation. The
interpreter will try to arrive at the correct interpretation of the subject, based only on the
interpreter’s impressive knowledge of facts about the subject. Are the facts available to
the interpreter in this scenario sufficient to determine a singular correct interpretation of
the subject?

At this point, a realist might diverge from interpretationists (and other realists). A
realist might claim that all the relevant facts about the subject suffice to determine a
correct interpretation of the subject. Interpretationists (and other realists) will insist that
multiple, equally correct interpretations remain. I will call the realists who diverge at this
point Naïve Determinationists.

This view should be considered naïve for reasons that are familiar within the
philosophy of science. The Naïve Determinationist denies the underdetermination of
theory by data. The underdetermination thesis holds that, even once all relevant data has
been considered, there will always be multiple theories that can account for all of it. One can choose any specific claim, and, in the absence of appeal to pragmatic constraints, one can always make adjustments to the theory to overcome apparent evidence against the claim. This could be accomplished, for example, by positing theoretical entities the existence of which would explain away the problematic evidence. Furthermore, two accounts might differ only in the manner in which they individuate objects; such accounts will make the same predictions, but will be distinct nonetheless. To be sure, some of the theoretical options will be needlessly complex, or practically useless, but our idealized interpreter is not heeding such practical considerations in the scenario we are currently considering. So, our interpreter has no way to resolve underdetermination. The data alone---even in an idealized scenario where there are no gaps in the data---simply do not entail one specific account and eliminate all the others. Thus, appeal to something beyond data---pragmatic factors---is required to identify one account as correct.

The same will hold for the various overall interpretations in the scenario we are currently considering. The data alone will not justify one choice of overall interpretation in lieu of all the rest. And the reasons for this do not depend on anything specific to thought or interpretation; they are reasons that apply to any attempt to account for a phenomenon. So, the reasons to believe in underdetermination are reasons to reject one form of realism, Naïve Determinationism. We should affirm, as interpretationists must

---

36 See Quine (1961).
37 See Quine (1969) and Davidson (1979).
38 I do not know whether anyone actually endorses Naïve Determinationism. Typically, complaints about those who promote underdeterminationism are directed against the importance the doctrine is taken to have for scientific methodology, and not against the doctrine itself. See Kitcher (1993), Laudan (1990), Laudan and Leplin (1991).
and most realists would, that there is more than one equally justified overall interpretation of the subject, insofar as one pays attention only to the data about the subject. And, in a context in which we have abstracted away from epistemic limitations, that shows that the facts about the subject do not metaphysically determine a lone correct interpretation.

Second Choice Point: Underdetermination and Scientific Pragmatic Factors

Having eliminated Naïve Determinationism from consideration, we arrive at our second choice point. Consider the same idealized scenario as before, but now allow the interpreter to appeal to general scientific pragmatic factors along with data. So, considerations about an overall interpretation’s simplicity or ease of use in making predictions can now be used to distinguish between different interpretations that are all capable of accounting for the data. Are the facts available to the interpreter in this scenario, along with the scientific pragmatic factors, sufficient to determine a singular correct interpretation of the subject?

A person who answers in the affirmative commits herself to realism. Because the data available to the interpreter have been supplemented by only general scientific pragmatic factors, and not factors that are specifically related to interpretive practice, an interpretationist must still deny the possibility of complete metaphysical determination in this scenario. Thus, according to the interpretationist, there should still be multiple,

---

39 i.e. not specific to the scientific study of thought
equally correct interpretations given complete data and general scientific pragmatic constraints as the only parameters in which the interpreter is to work.

The realist view that diverges from interpretationism at this point is more reasonable than Naïve Determinationism. To illustrate this superiority, consider the familiar case of the Copernican revolution. The heliocentric model of the solar system accounted for all the observations of heavenly bodies, as did the geocentric model together with its vast system of epicycles. Presumably, no future observation could have swayed a thoroughly committed geocentrist; further posits could always be made to account for any apparent failure of the model. But our justification for believing in geocentrism is not equal to our justification in believing heliocentrism. Either theory could be extended to account for all the relevant observations, but the geocentric theory proved to be cumbersome, ugly, and in need of frequent revision, whereas the heliocentric theory has none of these practical flaws. So, our justification for being heliocentrist far outweighs our justification for being geocentrist, and this is due to the kind of general scientific practical factors that the current realist view acknowledges and the Naïve Determinationist incorrectly deems unnecessary. In contrast to the thesis of Naïve Determinationism, there is nothing prima facie implausible about a realist’s claim that only one overall interpretation will be correct once general scientific pragmatic considerations are considered along with all of the relevant data.

Call a realist who holds this position a Scientific Realist about the thinking mind: like the phenomena studied by all sciences, thought attributions are subject to underdetermination even when we idealize so that all relevant data are possessed, yet
pragmatic standards such as parsimony and elegance can be used to privilege one assignment as the correct one.\footnote{See Fodor (1975), Block (1986), Antony (1994), Dretske (1995).} And, for a Scientific Realist, all relevant factors---data and pragmatic features---have been considered in the most recent scenario I have described. So, for a realist, the choice seems clear: she should deny that there can be more than one equally well justified interpretation now that all relevant data have been considered along with scientific pragmatic features, and conclude that the most justified interpretation is not merely an epistemic superior to the others, but that it is the metaphysically determined truth.

In contrast, interpretationist views, including my view, will continue to affirm the existence of multiple equally well justified interpretations even after general scientific pragmatic features have been considered. In support, we can appeal to the application of standards of rationality in the interpretive process. Since the goal with thought attributions is to rationalize the behaviors of the subject, there will always be competing ideas about what counts as the best rationalization, pragmatic features notwithstanding.

There are at least two reasons for this. The first of these arises from the fact that ideal rationality is not a perfect model for the cognitive and behavioral activities of subjects. When we apply a standard of rationality to a subject, we idealize away from three potential sources of conflict with our model: good reasoning based on typical desires and a factual error; poor reasoning based on typical desires and accurate beliefs; and good reasoning based on bizarre desires and accurate beliefs. Because we idealize away from factual error, poor reasoning, and bizarre desires all at once, we will always
have a choice between at least these three strategies for interpreting the oddly-behaving subject, one corresponding to each of the three co-idealizations. Thus, we require more here than the application of general scientific pragmatic constraints in order to anoint the result of any one of these three strategies as the correct interpretation. We do not merely need a standard against which to measure the data; we know from the beginning that subjects will fall short of the standard of ideal rationality. So, we also need an account of the source of failures to meet the standard, and the triune nature of the potential sources away from which we idealize means further interpretation is needed in order to give such an account. Even idealized interpreters could find that they face different social pressures in their quest to accomplish the practical, prescriptive goal of interpretation---social coordination through reinforcement of standards of rational social living---that gives the various interpreters reason to opt for different accounts of the subject’s error.

A second factor that prevents one interpretation from standing alone as the best rationalization of a subject’s activities is that interpreters can differ legitimately in their general assessment of what counts as best accomplishing the goals of interpretation. Determination of which assignment most successfully portrays the subject as rational will depend on the interpreter’s conception of ideal rationality, and there is no right answer with respect to what ideal rationality consists in. We might agree that it involves maximization of expected utility, but there is no fact of the matter as to whether one should prefer, for example, a maximin choice strategy or a riskier strategy that could deliver a greater reward. Interpreters are free to differ with respect to their views as to
which strategy is best, and different choices will lead to different ideas of which interpretation most satisfactorily portrays the subject as rational.

The fact that these choices regarding the precise nature of ideal rationality, and regarding the placement of blame in cases of defective activities, are actually available to interpreters, and would be available even under the idealized circumstances we have imagined, is sufficient to affirm that multiple equally well justified interpretations are available even given full data access and pragmatic considerations. In contrast, a Scientific Realist about the thinking mind must insist that, under such circumstances, a solitary correct interpretation will be identified. So, Scientific Realism should be rejected.

Third Choice Point: Dogmatic Realism vs. Interpretationism

A character we might call the Dogmatic Realist concedes to the interpretationist that there will be more than one equally well justified interpretation of a subject, even when the idealized interpreter has been given all the relevant data about the subject and is allowed to appeal to general scientific pragmatic considerations. But as a realist, the Dogmatic Realist still insists (for some reason) that among these equally well justified interpretations, there is just one that is correct. I call this view dogmatic because I see no possibility of offering support for it. The Dogmatic Realist would need to offer some reason to believe that the thinking mind must be the way exactly one of the equally well justified interpretations depicts it. Notice that the Dogmatic Realist must offer this
reason without appealing to any facts about the subject or any general scientific pragmatic factors. The Dogmatic Realist has already agreed that such features do not enable the idealized interpreter to distinguish one interpretation as superior to the others, and the idealization guarantees that such failure will occur only because of a genuine metaphysical indeterminacy, and not because of an epistemic shortcoming of the interpreter. Also notice that because the Dogmatic Realist is a realist, she is not allowed to appeal to the subject’s relations to interpretive practice. What is left for a Dogmatic Realist to appeal to?

I can only think of two strategies for the Dogmatic Realist:

(a) The Conceptual Strategy: argue that our concept of the thinking mind implies or at least suggests that minds must be determinate in the way the Dogmatic Realist claims.

(b) The Phenomenal Strategy: argue that, when we act as interpreters, it seems to us that we are discovering independently determined facts about the subject’s mind, and not contributing to the determination of facts about the subject’s mind.

Arguments involving analogous strategies might be given about the concept of number and mathematical practices. Consideration of these arguments will help clarify the nature of the Dogmatic Realist’s problems.

The mathematical analog of the Conceptual Strategy would be:

The concept of number seems to apply to an eternal, unchanging abstract order. If so, the status of this order is not impacted in any way by our justificatory tests. So, cases in which the justificatory tests fail to yield a determinate answer (even when these cases involve idealization) give us no reason to believe that the metaphysics of numbers is partially
determined by, and incompletely determined without, some contribution from mathematical practices.

And the mathematical analog of the Phenomenal Strategy would be:

When mathematicians make progress, it seems that they are discovering independently determined facts about the numerical order, and not contributing to the determination of facts about the numerical order.

Either of these arguments might be taken as reason to believe that the metaphysical phenomenon of truth can outrun the epistemic phenomenon of justification in mathematical discourse. And in the realm of mathematics, such arguments can be taken seriously. But parallel arguments in the realm of thought are not as plausible.

Consider the Conceptual Strategy. Unlike the concept of number, the concept of mind does not seem to apply to an eternal, unchanging order. So, unlike numbers, minds are not inherently insulated against the possibility that something external might play a role in determining their properties. That means that cases in which justificatory tests fail to yield a determinate answer as to which interpretation is correct cannot just be disregarded as revealing merely epistemic limitations, especially when we idealize to allow for all relevant facts to be accessible to the interpreter. Some reason must be given why we should believe that the thinking mind is determinate independently of its relationship to the interpretive process despite the principled inability of evidence and pragmatic principles to show us that it is. And the Dogmatic Realist has granted that there is such a principled inability. But there is no obvious reason to be offered, as there is in the case of mathematics.
Perhaps first-person access to the contents of our own mental states could be offered by the Dogmatic Realist as a reason to believe that the truths of mental content outrun the justificatory tests to which content assignments can be subjected. One could respond to the Dogmatic Realist by pointing to phenomena such as self-deception, or doctrines such as externalism about content, in order to call into question the claim that we have first-person access to the contents of our own mental states. But there is a more important move to make here. One ought to reply by pointing out that, in the idealized scenario wherein the interpretations are based upon all relevant facts about the subject, the subject’s belief that he knows the contents of his own mental states—the subject’s self-interpretation—would have been taken into account by the competing content assignments already. It is not an extra data point beyond the ones that the Dogmatic Realist has already agreed do not settle the epistemic issue between the various assignments. So it will not serve as an acceptable reason to support Dogmatic Realism here, and cannot bolster the Conceptual Strategy.

A Dogmatic Realist who opts for the Phenomenal Strategy holds that when interpreters make progress, it seems that they are discovering independently determined facts about the subject’s mind, and not contributing to the determination of facts about the subject’s mind. To assess this claim, I am going to split it into its conjuncts.

Conjunct one: “When interpreters make progress, it seems that they are discovering independently determined facts about the subject’s mind.” First, I think it is important to stress the difference between this claim and the weaker claim, “When

---

interpreters make progress, it seems that they are discovering facts about the subject’s mind.” The weaker claim is certainly true from an everyday vantage point. Even if we remove the hedge “it seems”, it takes a great deal of sophisticated theory to doubt the resulting claim, “When interpreters make progress they are discovering facts about the subject’s mind.” Someone in the grip of a theory might believe that there are no minds and so no facts about them for interpreters to discover (eliminative materialism), or that there are minds but interpretation is incapable of yielding facts about them (a thoroughgoing mysticism about the nature of the thinking mind), or that interpretation does not aim to reveal facts about minds but about commitments and entitlements to assert various claims (non-cognitivism applied to cognitive discourse). Even those who believe such things might still acknowledge that it seems like interpreters discover facts about subjects’ minds. So, I am not out to challenge that.

What about the claim that when interpreters make progress, it seems that they are discovering independently determined facts about the subject’s mind? Though I do not wish to deny this claim outright, I do want to point out that it can be doubted much more easily than its weaker cousin that does not contain the phrase “independently determined”. An everyday vantage point will probably regard it as true; from the everyday vantage point, one can imagine a member of a community of interpreters who is at the moment alone in nature, taking in her surroundings, and conclude that this subject would think the same things about her surroundings even if there were no interpretive practices at all.
But in this case, Socratic questioning easily displaces one from the everyday vantage point. Imagine again our practiced interpreter who at the moment happens to be alone with her thoughts about her natural surroundings. Now alter the case so that she is in the same surroundings, but there are no interpretive practices at all. The everyday intuitions lose force as we consider different ways in which our subject could have come to be alone in nature.

Are we to imagine that the subject was at one time a member of a community of interpreters? If so, then when we imagine her having the same thoughts despite an utter lack of interpretive practices, we are imaging her as a last-Crusoe-on-Earth. But in that case, perhaps it is merely her connection to past interpretive practices that leads us to believe she could have the same thoughts without them. So, our ability to imagine the last-Crusoe-on-Earth having the same thoughts as an otherwise identical being that is embedded in an interpreting community might not indicate the independence of facts about a subject’s mind from interpretive practice.

So, suppose we alter our imagined subject’s case so that, when we imagine her taking in her surroundings in a world with no interpretive practices, we are supposing that she was never a member of a community of interpreters. Should we still feel confident in asserting that she would have the same thoughts without any history with interpretive practice? I think we should not---subjected to these inquiries about the case it is imagining, common sense should realize that it has little to offer on the matter. Because our intuitions do not run particularly deep regarding these matters, we should be
unwilling to assert that interpreters seem to discover independently determined facts about subjects’ minds.

This thought experiment also casts doubt upon the second conjunct of the Dogmatic Realist’s premise: “When interpreters make progress, it seems that they are not contributing to the determination of facts about the subject’s mind.” I would not go so far as to say that it seems as though interpreters are contributing to the determination of facts about a subject’s mind. But neither do I think that it seems they are not. Instead, with respect to this issue, there is just not much phenomenology to speak of. When we consider the epistemology of the situation—remembering that the Dogmatic Realist has granted that equally well-justified assignments remain when all facts about the subject and pragmatic principles have been considered—-it seems reasonable enough that a subject’s mind might be metaphysically indeterminate when considered in abstraction from interpretive practice. There is no appearance to the contrary to rule this out, or even cast strong suspicion on it, once one has made the Dogmatic Realist’s concessions.

So, there is no good reason to accept the Dogmatic Realist’s premise “When interpreters make progress, it seems that they are discovering independently determined facts about the subject’s mind.”

It might be suggested that too much attention has been paid here to the phenomenology of the interpreter, and not enough to the phenomenology of the interpretee. To the interpretee, it does not seem that his own thoughts await constitution by the activity of an interpreter, and surely this ought to count in favor of realism. But this phenomenology on the part of the interpretee is neither surprising nor troubling to an
interpretationist. The act whereby a person turns his attention to his own thoughts, to see whether they are fully determinate or not, places him in the role of interpreter as well as interpretee. It involves an act of self-interpretation. Thus, a case in which we are asked to imagine the phenomenology of the interpretee when he attends to his own thoughts is not a case in which the role of the interpreter goes unfilled. The reflective interpretee finds his own thoughts always fully constituted because he himself is doing the constituting in every such instance.

At this point, we have reached a limb of our decision tree from which no Realist branches grow. So, I am now in a position to deny that a single overall assignment of propositional attitudes can embody an interpreter’s discovery of the independently determined set of mental states for the interpretee. Remember, though, that I do not want to deny: “When interpreters make progress, it seems that they are discovering facts about the subject’s mind.” Indeed, I am willing to assert the claim even without the hedge. So, I accept “When interpreters make progress they are discovering facts about the subject’s mind” but deny that an interpretation can embody an interpreter’s discovery of the independently determined set of mental states for the interpretee. So, what facts about the subject’s mind is an interpreter discovering, given that the interpreter is not discovering the truth about which assignment uniquely and correctly specifies the subject’s independently determined state of mind?

The short answer is: the interpreter is discovering facts that limit the range of acceptable content assignments to the subject, by discovering which overall identifications of pattern vs. noise will adequately cover the facts about the subject, given
the constraints set by pragmatic principles. This short answer makes reference to the notions of pattern and noise, notions that are employed by the interpretationist Daniel Dennett in his explanation of his view of the thinking mind. So, I will elaborate on the short answer by turning to a discussion of Dennett’s scientific interpretationism, and the manner in which the notions of pattern and noise can be exploited by the theorist of thought. We will then be at our final choice point: Scientific Interpretationism vs. Non-scientific Interpretationism.

Pattern, Noise, and Interpretation as Discovery of Facts about Minds

Dennett holds that any interpretation must treat some data as fitting a pattern and other data as noise. For Dennett, the interpreters in question are scientists, so they are interested in giving elegant accounts that fit systematically with the body of science, explain the data in question, and yield accurate predictions. But trade-offs will typically have to be made between these desiderata, so it will be up to the interpreter to make them by selecting one assignment of pattern-to-noise from among those that are equally justified given the data and pragmatic considerations.

What is it for an interpreter to treat some data as fitting a pattern and other data as noise? A pattern exists whenever a collection of data points can be represented more elegantly than by simply listing it point by point. In the case of intentional interpretation, the data are human activities and the interpreter tries to represent them more elegantly by

---

fitting them, and the cognitive processes that cause them, to rational patterns. Rational patterns are distinguished among the class of patterns by their propositional/inferential structure and characteristic norms, which include consistency of beliefs, transitivity of preference rankings, inferential soundness, attentiveness to one’s reasons, conformity of will to reason, and maximization of expected utility.43

To take a simple example, suppose my data concerns the behaviors of a subject who has written the following symbols: “1*1=1; 2*2=4; 3*3=9…9*9=81.” I could represent these data by listing each behavior: the subject drew a character of shape <1>; the subject drew a character of shape <*> to its right; and so on. But it is much more elegant to simply say that the subject squared the single-digit positive numbers. I have found a way to represent the subject’s behavior as fitting a pattern.

In this case, the behaviors fit the pattern perfectly, so there is no noise. But that is rarely the case. Typically, there will be some recalcitrant data. Suppose a subject has written a long list of symbols easily interpretable as addition problems. But the list contains one oddity: at one point, the symbols “68+57=5” appear. An interpreter who treats the subject as performing addition problems will treat the majority of the symbols as pattern and the unusual string as noise---data that does not fit the pattern.

Of course, if the interpreter has just finished reading Kripke’s *Wittgenstein on Rules and Private Language*,44 the interpreter might regard the subject as performing quaddition problems. The quaddition function works exactly like addition as long as each of its arguments takes a value less than 57, but yields a “quum” of 5 whenever an
argument takes a value 57 or greater. An interpreter who took the subject to be quadding would treat “68+57=5” as pattern. If the list contained only one other problem involving argument values 57 or greater, say “70+60=130”, then that string would be treated as noise.

With a little more ingenuity, one could craft a rule that yielded no errors at all on the part of the subject—say, schwaddition, which is similar to quaddition except that it requires exactly one argument to the function to take a value between 56 and 61 in order to yield 5 rather than the sum. A trade-off for rendering the subject error-free in such a way is that this rule itself is less elegant. The inevitability of this type of trade-off largely motivated the decision to occupy the Interpretationist-Dogmatic Realist limb of our decision tree, and Dennettian talk of patterns and noise captures it well.

Hopefully, the examples of addition, quaddition, and schwaddition serve to illustrate another point about patterns: their ubiquity. Patterns are remarkably easy to come by; for any bit of data, one can generate indefinitely many patterns that it fits. The patterns themselves all exist, regardless of whether it is useful to attend to them. Consider again the case in which our subject generated many math problems, and made either one error in addition or one error in quaddition. Though we might have pragmatic reasons for choosing addition rather than quaddition for use in a proper interpretation of our subject, the quus-function’s pattern exists in the data every bit as much as the plus-function’s pattern does. It is a fact about the subject’s behavior that it contained each of those patterns to an equal extent. So, it is a fact about the subject’s mind that it is as readily interpretable, on the basis of these data, as containing intentions to add as it is
interpretable as containing intentions to quad. And it is this kind of fact about a subject’s mind that is waiting, fully constituted, for discovery during the interpretive process.

This has been the more elaborate answer to the question, “What facts about the subject’s mind is an interpreter discovering, given that the interpreter is not discovering the truth about which assignment uniquely and correctly specifies the subject’s independently determined state of mind?” Remember that the short answer was: the interpreter is discovering facts that limit the range of acceptable content assignments to the subject, by discovering which overall identifications of pattern vs. noise will adequately cover the facts about the subject, given the constraints set by pragmatic principles. So, interpretationists have a way to account for the fact that interpretation involves discovery of interpretation-independent facts about a subject’s mind while preserving the claim that interpretive practice plays a constitutive role in the fixing of facts about a subject’s mind.

Final Choice Point: Scientific or Non-scientific Interpretationism

This brings us to the final branching of our decision tree. Our options are Scientific and Non-Scientific Interpretationism, and we must determine which one of them captures the truth about the thinking mind. Ultimately, I will side with non-scientific interpretationism. Scientific interpretationism does not fully acknowledge the authoritative status that everyday practice holds with respect to the nature and existence of thought.
For the Dennettian, thought is a natural kind. Granted, it is a natural kind that requires some choices to be made by the scientist acting as interpreter; that is why a Dennettian is a scientific interpretationist. But it is nonetheless within the purview of science; that is why a Dennettian is a scientific interpretationist. Because the Dennettian is first and foremost interested in the scientific study of the thinking mind, if a case arose in which we had to choose between satisfying the desiderata of science or the desiderata of everyday practice, the Dennettian would side with science.45

A non-scientific interpretationist would side with everyday practice. For the non-scientific interpretationist, thought is a social kind, and does not require any validation from science. The dispute is about the locus of the final authority over the existence and nature of thought. The Dennettian chooses the scientific study of cognition; the non-scientific interpretationist chooses everyday practice.

So, if one wants to argue that we should accept non-scientific interpretationism over scientific interpretationism, one must show that everyday practice has the authority to establish the existence of thought regardless of what science determines about cognition’s structure. At this point, we once again engage a debate that was only tentatively settled in Chapter 1. There, we asked whether “thought” attempted to name a natural kind or a social kind. I pointed out there that the everyday practice in which cognition is regarded as having the structure of thought is subject to normative pressures that scientific practice is not. And I argued that this divergence in purposes shows that

45 See Dennett (1989).
the everyday notion of thought does not depend for its legitimacy on whether or not it best serves scientific purposes.

But one might recognize the legitimacy of the notion of thought for social purposes without thereby acknowledging that thought exists. Regardless of our need to treat the vehicles of cognition as propositional attitudes during everyday interpretive practice, are the vehicles of cognition really propositional attitudes? In Chapter 1, I gave a short answer that relied heavily on folk intuitions about what really exists. According to those intuitions, jade exists and dinosaurs existed even though neither “jade” nor “dinosaur” turned out to name a natural kind; by analogy, the vehicles of cognition can be propositional attitudes even if “thought” turns out not to name a natural kind. With the Dennettian notions of pattern and noise at my disposal, I can now give a longer answer. This longer answer will show that “thought” names a class of entities whose existence as such can be established regardless of whether it names a natural kind. I will accomplish this by showing that science, even idealized science, is not always the final authority on what exists and need not be privileged over everyday practice with respect to the nature of cognition. Thus, it will allow us to choose non-scientific interpretationism as the correct approach to the thinking mind.

As a preliminary, I want to acknowledge the appeal of the Dennettian approach. Scientific practices seem to have earned the right to be considered authoritative with respect to the ontology of the empirical world. Science has a long history of vastly improving our epistemic state with respect to the world around us. Science has

---

supplemented simple folk observations with a rigorous methodology that treats data systematically and often yields elegant, cohesive accounts of diverse phenomena. In cases where there has been conflict between the folk wisdom on a topic and the conclusions drawn by science on the same subject, the claims of science have repeatedly triumphed when put to the tests of prediction and application to practical purposes. And though any contemporary scientific practice will diverge from the ideal, and so stand some chance of becoming outdated as new evidence is gathered and new theories are inspired, sciences do often reach points at which it is clear enough that an ideal practice will have to somehow include certain notions. In such cases, it seems that science should be treated as authoritative with respect to the existence of things that fall under that notion, and the non-existence of things that fall under competing notions. So why should it be any different when the topic is cognition?

As long as we are merely interested in giving the best systematic, descriptive account of the nature of cognition, it should not be any different. Science should indeed be treated as authoritative about the way that cognitive processing is accomplished by cognitive systems. And at some point, science will be sufficiently well-informed to justifiably forecast which structure an ideal science would find in cognition, the structure of propositional attitudes or some other structure.

This acknowledgment may seem like doom for a non-scientific interpretationist. But, as I have already discussed, the everyday practice of thought attribution is much more than an attempt to systematically describe the nature of cognitive processing. It serves normative purposes vital to the flourishing of human societies, and it can only
serve these purposes by presupposing that cognition takes the form of propositional attitudes, a presupposition with which science might not concur. The possibility that cognitive science and everyday interpretive practice are working at cross purposes, together with the immense importance of the everyday purposes, give us reason to pause and more cautiously consider whether science should be granted full ontological authority with respect to the structure of cognition.

This does not, however, suffice to establish that everyday practice has the authority to establish the existence of thought independently of the verdict of science. If there comes to be a scientific consensus that an idealized scientific practice would not recognize cognition as having the structure of thought, one available response would be that science is ontologically authoritative with respect to the empirical world, and so propositional attitudes are merely an indispensably useful fiction.47

The Dennettian scientific interpretationist is supposed to be protected from this potential slide to eliminativism by the existence of real patterns. Everyday practice works for a reason, and the reason is that it identifies real patterns. Since we already know the patterns exist, there is no danger that science will dismiss them, and with them the idea that cognitive processes have the structure of thought. To do so, science would have to dismiss the existence of patterns that are shown to exist by the success of everyday practice. But science will not do the absurd, so there is no threat of science deeming thought non-existent.

The problem with that line of thought is that there is a chance that science will dismiss the existence of thought, even though it is based on real patterns, without science thereby doing the absurd. Real patterns are easy to come by. Everyone should believe that there are patterns that correspond in a rough, noisy way to propositional inference in the cognitive processes of linguistic humans. If there were no patterns that even came close to propositional inference in the cognitive processes of linguistic humans, interpretation would not have much instrumental value. Nonetheless, science might reasonably deem thought non-existent if another conception of cognition yields a better pattern-to-noise ratio, and so fulfills the purposes of cognitive science more efficiently than the everyday notion of thought.

So, the existence of real patterns provides no protection against a possible slide from scientific interpretationism to eliminativism. Any view that treats science as authoritative about the structure of cognition will be vulnerable to a similar slide. But real patterns can be employed by the non-scientific interpretationist to help make the case that science should not be treated as the sole authority with respect to the real structure of cognition. For, when we consider that there are many real patterns to be found in the cognitive processes of linguistic humans, we can see that the truth about cognition might be this: there is no such a thing as the structure of cognition. As is the case with any phenomenon, there will be many ways to break cognitive processes down into pattern and noise. Because each of these identifies real patterns in cognition, each of these is a prima facie candidate to count as the structure of cognition. Most of these will not be particularly well-suited for any purpose, and the vast majority will not be the best-suited
for any important, systematic purpose. But some will be. And in order to declare any of them the structure of cognition, it would have to be best-suited for every important, systematic purpose we have in attending to the structure of cognition.

This is where cognition differs from most phenomena that human beings wish to confront systematically. For most phenomena, our important purposes in confronting them systematically run no risk of coming into conflict with whatever assignment of pattern-to-noise serves science best. For most phenomena, our needs in systematically confronting them are exhausted by a desire to efficiently predict and explain them, and that is exactly what science is set up to accomplish. As a result, whichever assignment of pattern-to-noise best serves scientific purposes with respect to that phenomenon will stand as the only legitimate candidate to be the structure of that phenomenon.

But we have a further, normative need for a systematic account of the structure of cognition. The propositional-inferential structure of cognition is presumed by some of our most important social practices in order that we may judge and enjoin according to rational standards. To treat a subject’s actions as the results of everyday thought is not just to find some pattern that the action approximates. It is also to choose to assess the subject’s actions according to some norms and not others. Interpreters are arbitrators as well as discoverers. The fact that interpretation always calls for some decision on the part of an interpreter to treat some norms and not others as in governance has important results. It provides a principled barrier to all attempts to reduce rule-following to mere pattern-embodying. This means that the process of interpretation should not be treated as
a merely descriptive project, a mere discovery of patterns, or even a mere selection of patterns that best serve scientific, descriptive purposes.

In the case of cognition, no matter how useful a particular pattern might be for descriptive, scientific purposes, we have an important normative reason to confront a phenomenon systematically in ways that might come into conflict with science. Nonetheless, we are constrained in our choice of norms by real patterns; we are not merely projecting structure that is not there onto cognitive processes. And so, in the case of cognition, there could be more than one legitimate contender for the title of the structure of the phenomenon. And this means that we have good reason to believe that there might be no such thing as the unique structure of cognition. Science and everyday practice might both have the authority to find a structure in cognition. Even though the structures they find might be distinct, there might be no inconsistency in asserting that cognition has each structure.

With this combination of factors in place, I see no further reason to deny the authority of everyday practice to determine that thought exists, independently of the outcome of science’s search for whichever structure of cognition best serves scientific purposes. This reinforces the anti-eliminativist conclusion drawn in Chapter 1. And it supports non-scientific interpretationism over scientific interpretationism. Because non-scientific interpretationism considers the everyday treatment of cognition qua thought to be legitimized independently of science’s verdict on the nature of cognition, it is free to accept the possibility that cognition has a dual structure in a way that the Dennettian is not. Dennettians accept the reality of multiple, divergent assignments of pattern-to-noise
with respect to a given body of data. But they must privilege some patterns over others, and as scientific interpretationists, Dennettians must allow the descriptive purposes of science to have the final authority on which patterns are privileged. As far as the non-scientific interpretationist is concerned, however, cognition can be viewed with equal legitimacy as having the structure of propositional attitudes and as having whatever structure science attributes to it, and this is as it should be on a view of thought developed in recognition of the existence of real patterns.

So, everyday interpretive practice has the legitimate authority to determine that a specific structure, the structure of propositional attitudes and inferences, is present in at least some cognitive processes of fully linguistic humans. Everyday practice has this authority only because of the confluence of many key facts. The fact that our everyday practice assigns a structure to all cognitive processing it encounters shows that it confronts the phenomenon of cognition systematically. The fact that the confrontation is systematic allows us to consider the everyday practice as a candidate to be authoritative on the structure of cognition. Furthermore, the everyday practice serves important normative purposes that can only be served by a practice that treats cognition as having the structure of propositional inference. The fact that the practice serves normative purposes shows that it is constrained by pressures that diverge from those of science. This opens the possibility that it will yield a verdict with respect to the structure of cognition that differs from that of science. The fact that important purposes can be served only by making an ontological presupposition that might conflict with science gives us reason to question whether science is the only legitimate ontological authority
here. Additionally, the success of our everyday practice establishes the *existence of real patterns* that approximate thought in the cognitive processes of (at least) linguistic humans. Contrary to the hopes of the scientific interpretationist, the existence of these patterns does not guarantee the scientific recognition of propositional attitudes as real. The existence of real patterns does, however, provide the non-scientific interpretationist with a response to the eliminativist’s claim that propositional attitudes are merely a convenient fiction projected onto linguistic human cognitive processing.

The final step required to justify the claim that everyday practice has the requisite ontological authority to establish the existence of thought was the realization that this does not rob science of the authority to establish that cognition has some other structure. An apparent dispute between everyday practice’s claim that cognition has the structure of thought and an ideal cognitive science’s possible claim that cognition has some other structure turns out to be merely apparent once we consider the ubiquity of real patterns. Each practice could be discerning in the same data an overall assignment of pattern-to-noise that best serves its purposes. The claim that cognition has structure A is not necessarily inconsistent with the claim that cognition has structure B. So, the claim that everyday practice has the authority to establish that cognition has the structure of thought does not rob science of the authority to establish that cognition has some other structure as well. Thus, the notion of real patterns allows us to recognize multiple structures as present within the same phenomenon, and thus allows us to recognize non-scientific practices as capable of determining that a phenomenon has some structure or other without challenging the authority of science over the nature of the empirical world.
I conclude that non-scientific interpretationism is the correct approach to the thinking mind.

Conclusion

When we inspect everyday interpretation, we are confronted with a practice that succeeds in serving a vital normative purpose by choosing to focus on one particular set of real patterns present in human cognition. It serves this purpose regardless of whether focusing on that set of patterns best serves the descriptive purposes of science. Yet because there can be a multiplicity of real patterns within the same phenomenon, it is no threat to the verdict an ideal science might deliver regarding cognition’s structure, removing our motivation for denying everyday interpretation an authoritative status with respect to cognitive structure. An ideal science would focus on whichever structure of cognition is most useful for scientific purposes. Everyday practice focuses on a structure of cognition—-the structure of propositional attitudes and inferences—-that is most useful for certain social purposes—-the coordination of rational social living and the reinforcement of the standards that govern it.

For these reasons, we can conclude that our everyday term “thought” refers to a social kind, a kind that might not be coextensive with any natural kind. Applying this conclusion to the search for the best philosophical account of the thinking mind, we can rule out any views on which “thought” is treated as attempting to name a natural kind, whether successfully or not. This eliminates many realist views, and scientific
interpretationism, from contention, as well as eliminativism. The remaining realist views are either dogmatic or implausibly deny the underdetermination of theory by data. Non-scientific interpretationism alone survives as a plausible, well-motivated view of the thinking mind.

In Chapter 2, I have invoked the notion of a pattern in order to explain one requirement a subject must meet in order to be correctly interpreted as a thinker: a rational pattern must be present in the subject’s cognitive and behavioral activities. But interpreters do not merely find rational patterns in subjects’ activities; they also judge subjects according to rational standards. And this means that subjects must meet requirements for being appropriately held to those standards in order to count as thinkers. Influential interpretationists\(^48\) have argued that non-linguistic beings cannot meet these requirements. In Chapter 3, I will survey the purported requirements and the ways in which they can be used to deny thought to non-linguistic beings. In Chapter 4, I will show how non-linguistic beings can meet all of the legitimate requirements for rational assessment. Hopefully, the value of such a project will be apparent even to those who remain unconvinced that interpretationism is the correct view of thought; if successful, my account will show how to accommodate the most important insights of interpretationists, and meet the most important challenges that have been raised to the claim that non-linguistic beings can be thinkers.

\(^{48}\) See Davidson (1975 and 1986c), Brandom (1994), Malcolm (1972-3).
Chapter 3: From Requirements for Thought to Arguments for Linguistic Priority

Introduction

In Chapter 1, we saw that to treat cognitive processes as instances of thought is to treat them as having the structure of rational inference. Furthermore, when we treat cognitive processes as thought, we assess them according to the norms of rationality. In Chapter 2, we focused on the essential role played by the interpreter in the process of thought determination. Because the nature of rational assessment leaves open some matters for decision---what to treat as pattern and what to treat as noise when error of some kind must be attributed, and whether interpreting our subject as playing a maximin or maximax choice strategy would render it more rational, for example---rational assessment cannot occur without someone actively occupying the role of interpreter. This explains why interpretationism is the correct view about the thinking mind; the metaphysical constitution of any thought is partially determined by the active interpreter.

Chapter 3 also centers on the fact that thought is cognition assessed according to the norms of rationality. But our focus this time will be on the potential interprettee, and the requirements such a being must meet in order to be credited with thought. We already encountered one such constraint in Chapter 2: the behaviors and cognitive
activities of the interpretee must display a pattern that at least roughly embodies the structure of rational activity.

But to be a thinker, it is not enough that a potential interprettee’s cognition and behaviors fit a rational pattern. Interpretation does not merely find rational structure in a subject’s activities, but judges the subject according to rational standards, as well. Because of this, a subject must meet requirements for being appropriately held to such standards. Some of these requirements are generic ones that specify what is necessary in order for a being to be appropriately held to any normative standard. Others are requirements placed on potential thinkers in virtue of the fact that we are trying to judge them by the norms of rationality specifically.

These requirements for appropriate judgment by norms are central to the case that only language users can be thinkers. Some of the most influential interpretationists have argued that requirements of this kind can only be met by language-using subjects. And it is only by showing how a non-linguistic being could meet the requirements that one can establish the possibility of non-linguistic thinkers. In Chapter 3, I will survey the purported requirements and show how they have been—and can be—used to argue for linguistic priority. In Chapter 4, I will demonstrate that such arguments fail by showing that some of the purported requirements are not legitimate, and by showing how non-linguistic beings might meet the remaining requirements.

If we are only interested in seeing how closely a being’s behaviors come to meeting a standard, we can apply any standard we like without thereby assigning any

49 See Davidson (1975 and 1986c), Brandom (1994), Malcolm (1972-3).
kind of authority to the being. On the other hand, to treat a being as bound by a norm is to treat it as being responsible for its activities vis-à-vis that norm’s demands. So, when we treat a being as bound by a norm, we presume that the being’s behaviors that are relevant to its meeting the norm’s demands are within the being’s influence. When we find that the being in question has such authority, the results of our application of normative standards to its behavior can reveal the truth about normative statuses the being occupies. If, however, we find that a being cannot be treated as having this authority over its behaviors, it will be inappropriate to attempt to assign the being a normative status.

But what must a subject be like in order for an interpreter to find that the subject has such authority? What other requirements must a subject meet in order for us to legitimately hold it accountable to the norms of rationality? And why would someone think that this required the subject to use a public language? In the following section, I will provide a brief overview of seven purported requirements on holding a subject and its cognitive activities accountable to the norms of rationality: intensionality, extensionality, objectivity, reflection, control, institution of standards, and composition from specific concepts. The remainder of the chapter will fill in the details of how these constraints can be used to generate ten arguments for linguistic priority: The Davidsonian Holism, Network, Error, Concept Cluster, Generality, and Master Arguments; the Basic Institutional Argument; the Kripkensteinian Argument; and the Brandomian Syntactic Priority and Deliberation Arguments.
Purported Requirements a Subject Must Meet to Count as a Thinker

To treat a subject as having authority over a certain class of its behaviors, we must be able to recognize the being as meeting several requirements. At least seven proposed requirements are relevant to the issue of whether there can be non-linguistic thinkers. One might require a thinker to have control over the relevant behaviors. What exactly this amounts to is controversial. For one thing, there are debates about whether control is compatible with a deterministic view of our world. More importantly for our purposes, it is unclear exactly which behaviors must be within the being’s control in order to establish its authority in a given case. Is it enough that the being has control over some of its behaviors, as Donald Davidson holds, or must it have control over the very behaviors that the interpreter is considering?

Despite these complications, one can still make sense of which kinds of behaviors a requirement of control might discount from among those over which a being has authority. An involuntary response brought about directly by an environmental stimulus, for example, is not within a being’s control. The involuntary response might result in the being’s preferences being satisfied, and so be judged favorably when measured by a rational standard, yet that does not give us sufficient reason to hold the being accountable to rational norms when assessing that particular behavior. If the being is to have authority over such an automatic and uncontrollable response, it can only be because its

---

50 See Davidson (1975).
control over some of its behaviors confers authority over other of its behaviors---as in the Davidsonian line mentioned in the previous paragraph.

One might also see reflection on one’s thoughts as required for authority, either because it is required for control, or in its own right. One might hold that a being only has full control over behaviors for which that being explicitly considers different options and deliberately undertakes one of them. If one wants to use the term “will” to refer to the faculty by which beings control actions, and the activities of that faculty, then the claim would be that beings only will actions that have been deliberately undertaken, where deliberation is understood to be a reflective process. All other behaviors arise entirely from causal factors that are outside the attention of the being’s action-controlling faculty, and so are not controlled by the being. Or one might bypass the talk of control and simply hold that a being only attains full authority over behaviors by explicitly considering options and deliberately undertaking some of them. In either case, there will be room for debate here between a view on which a being merely needs to be able to deliberate in order to be granted authority over any, or many, of its behaviors, and a view on which the being must have deliberately chosen the very actions the interpreter is considering.

One might also argue that we should not hold a subject accountable to a norm unless it grasps certain concepts. This requirement of conceptual grasp is subsidiary to the reflection requirement. It insists that one must be able to think about one’s own thoughts as such in order to deliberately choose a course of action. So, one must have concepts such as THOUGHT, BELIEF, etc. in order to meet the reflection requirement.
A more general point about concepts is that each one applies to some items and not others. To treat a being as applying concepts, then, is to treat it as accountable not just to the standards of rationality, but to the standards for each concept’s application, as well. And we must treat a being as applying concepts in order to treat it as a thinker. As familiar examples involving co-referential terms like “Morning Star” and “Evening Star” show, reasoning cannot be judged rational or irrational unless one can specify how the subject thinks of the objects of its thoughts. So, rational assessment involves attribution of concepts to a subject, i.e. the concepts in terms of which the being thinks of the objects of its thoughts. Thus, another feature of thinkers is their accountability to standards for application of concepts.

The fact that reasoning can only be judged rational or irrational when one can specify how the subject thinks of the objects of its thoughts has another consequence. To even reach a point at which one can ask whether a subject is genuinely accountable to the standards for a particular concept’s application, one must have already been able to tentatively identify that concept as the one by which the subject thinks of the object of its thought. Thought attribution requires that we pick out one of the many ways in which a subject could think of an object as the way the subject actually does think of the object; this is an intensionality requirement.

The fact that we need to identify the way in which a thinker thinks of the objects of its thoughts also implies an extensionality requirement. The subject must be understandable as directing a thought toward an object in the first place. To treat a being
as bound by rational norms, one must be able to make sense of it as thinking about certain objects.

The intensionality and extensionality requirements together capture the representational dimension of thought. Beliefs represent the world as being objectively one way or another. This yields an objectivity requirement: to treat a being as bound by norms of rationality implies that it has, or is in, subjective states to which it is appropriate to apply the standard of objective truth. Davidson supplements the objectivity requirement with the first version of the conceptual grasp requirement; he treats grasp of the concept of belief as essential to having subjective states that function to represent an objective world.51

Finally, one might hold that there is an institutional requirement. The idea here is that a norm has no power to bind a subject apart from its institution in a practice, and a subject must itself be a practitioner in order to be bound by the norms that the practice institutes. One might hold that normative binding is contractual, and that only by contributing to the practice’s institution does a being give consent to be bound by its norms. Or one might hold that authority over one’s movements is a status earned only by having played particular roles or mastered certain skills within a practice. One might attempt to apply this requirement to the norms of rationality generally, or one might focus specifically on either the standards for concept application or those of truth. On any of these lines, the subject must actually participate in a practice in order to be bound by its norms.

51 See Davidson (1986c).
In summary, I have introduced seven purported requirements for attribution of thought to a subject. They are the requirements of:

- control
- reflection
- conceptual standards
- intensionality
- extensionality
- objectivity
- institution

Additionally, one should keep in mind that the reflection requirement can be seen as an elaboration on the control requirement, and that the requirements of reflection and objectivity can be supplemented by a requirement that the subject grasps the notion of thought, or the notions of various kinds of thought, such as belief. In the following sections, I will consider the ways in which these purported requirements can be used to argue against the possibility of non-linguistic thinkers.

From the Seven Requirements to a Requirement of Language Use

At this point, one would like to give a literature summary. One would like to identify and clearly explicate arguments in the literature from any of the seven requirements to the claim that a subject must be a language user to be a thinker. But, as Davidson pointed out in “Thought and Talk”, philosophers have preferred taking a stand
on the issue of linguistic priority to giving an argument for it.\textsuperscript{52} Unfortunately, this has not changed much in the years since “Thought and Talk”. We see Robert Brandom claiming to fill in a Davidsonian argument schema, proceeding to identify the schema,\textsuperscript{53} but never telling the reader how exactly he is filling it in. Saul Kripke’s Wittgenstein and Tyler Burge engage us in thought experiments that seem to imply that aspects of mental content depend on linguistic practice, but neither insists that thoughts can be attributed only to linguistic subjects.\textsuperscript{54} Even in Davidson’s own work, we see him “marshaling considerations” and “presenting these considerations as an argument.”\textsuperscript{55} rather than spelling out, premise by premise, exactly what the structure of such an argument is supposed to be. Perhaps this is because, as Davidson puts it, there are “strong, if turgid, currents”\textsuperscript{56} running beneath the intuitions that guide philosophers to their views on this issue. Very well then. We will need to brave the currents if we are to bring any clarity to these matters.\textsuperscript{57}

In this section, I will attempt to do just that. I will take the considerations that others have marshaled and fit them into ten arguments for linguistic priority. The seven purported requirements for thought that I have listed above represent some of the key considerations that have been marshaled against my view that there can be non-linguistic thinkers, and at least one of the seven requirements is exploited by each of the ten

\textsuperscript{52} See Davidson (1975), p. 156.
\textsuperscript{54} Kripke (1984) and Burge (1979b).
\textsuperscript{55} See Davidson (1986c), p. 478.
\textsuperscript{56} In Davidson (1975), p. 156.
\textsuperscript{57} For other statements of the Linguistic Priority Thesis, see the Introduction to Preston (1997).
arguments. The goal of this section is to show how the seven requirements can be exploited to support a requirement of language use.

I will be guided in this process by the literature on the topic. That literature exists for a reason, and I will do my best to give arguments that respect the intuitions that drove its creators. But since even they did not feel comfortable explaining how their own considerations fit together into arguments, I shall not claim to have unveiled the truth about how their arguments were intended to work. Thus, I will call the lines of argument that I present “Davidsonian”, “Brandomian”, etc. rather than identifying them as Davidson’s, or Brandom’s, etc. For my purposes, what is most important is to have a thorough list of the prima facie plausible arguments that one could give for linguistic priority based upon the considerations that have been marshaled, and not a list of the arguments that have been given.

Considerations that Davidson Marshals in “Thought and Talk”

Davidson’s “Thought and Talk” is one of the few papers to explicitly argue against the possibility of non-linguistic thinkers. Davidson’s thesis in that paper is “that a creature cannot have thoughts unless it is an interpreter of the speech of another.”58 So, we are promised an argument that only interpreters of language can be thinkers.

In many respects, Davidson agrees with the account of thought I am defending in this dissertation. We agree that thoughts have the same propositional structure as

sentences. And he holds that “… talk apparently of thoughts and sayings does belong to a familiar mode of explanation of human behaviour and must be considered an organized department of common sense.”\(^{59}\) So we also share a commitment to understanding thought in terms of the everyday practices in which thoughts are attributed. And we agree that this practice involves finding patterns in the activities of a subject: “Adverting to beliefs and desires is … a way of fitting an action into a pattern of behaviour made coherent by the theory.”\(^{60}\) Furthermore, we agree that the patterns in question are given shape by the norms of rationality: “A characteristic of teleological explanation … is the way in which it appeals to the concept of reason. The belief and desire that explain an action must be such that anyone who had that belief and desire would have a reason to act in that way … in light of the content of the belief and the object of the desire.”\(^{61}\)

I also acknowledge the existence of many of the parallels between thought contents and linguistic meanings that Davidson emphasizes. Everyday practice treats sentences as expressions of thoughts, so I agree with Davidson’s claim that “if we know [a subject] holds the sentence true and we know how to interpret it, then we can make a correct attribution of belief. Symmetrically, if we know what belief a sentence held true expresses, we know how to interpret it.”\(^{62}\) Because everyday practice treats language as an expression of thought, interpretation of a linguistic subject must involve fitting both thought contents and sentence meanings to a unified rational pattern. The interpretive

\(^{59}\) Ibid. p. 158.  
\(^{60}\) Ibid. p. 158.  
\(^{61}\) Ibid. p. 158. I would hesitate to give such an internalistic treatment of the notion of “having a reason”, but Davidson’s point is clear enough.  
\(^{62}\) Ibid. p. 162.
method, when applied to linguistic beings, must allow sentence meanings to match up to thought contents in the right ways.

The need for proper correspondence between the cognitive and the verbal blocks any attempt to give a complete account of either one independently of consideration of the other. Suppose one tried to first assign meanings to all of a subject’s sentences in a way that rendered the subject’s linguistic interactions reasonable, with the idea that thoughts could then be attributed by matching contents to meanings one by one. One would risk finding that the assignments of contents failed to render rational the subject’s overall pattern of linguistic and non-linguistic activity. Or vice versa: suppose one tried to work out the being’s thoughts first, perhaps using a technique for naturalizing mental content, and then tried to simply read off linguistic meanings from the account of thought content. One would then assign sentential meanings without regard for the way those sentences interacted with the sentences of the being’s interlocutors, potentially failing to make good sense of the being’s linguistic behaviors.

Therefore, I agree with Davidson that “the project of interpretation of linguistic beings [is] a matter of working out together meanings and thoughts.” When both language and thought are present in a subject, neither thought contents nor linguistic meanings can be totally explicated independently of the other. My disagreement with Davidson emerges only when we consider cases in which a subject is non-linguistic. Though the process of thought attribution cannot be completed for a linguistic subject

---

63 For example, see the account of Fodor (1990) or Dretske (1988).
64 Ibid. p. 162.
without assigning meanings to her sentences, in the case of non-linguistic subjects, there are no sentences to interpret.

One reaction to this fact would be to claim that the interpretive process cannot be completed for non-linguistic subjects, because one key element is missing. But this reaction is unreasonable given only the considerations I have mentioned so far. For all I have conceded so far, the only reason why thought attributions must be constrained by sentential meaning assignments is that an independently completed account of mental content might fail to make sense of the subject’s linguistic interactions. With a non-linguistic subject, there are no such interactions to make sense of, so there is no such worry. So, if one wishes to deny thought to non-linguistic beings, the fact that thought attributions cannot be worked out for linguistic beings without consideration of their linguistic activities will not alone suffice. Indeed, one might have the opposite reaction, and claim that the interpretive process can be completed much more easily for non-linguistic beings, simply because we need not worry about simultaneously working out meaning assignments and thought attributions.

But this idea leads us to the first of Davidson’s serious worries about attributing thoughts to non-linguistic beings. He points out that “where one constellation of beliefs and desires will rationalize an action, it is always possible to find a quite different constellation that will do as well. Even a generous sample of actions threatens to leave open an unacceptably large number of alternative explanations.”\(^{65}\) The issue here is the holism of the mental: the fact that beliefs and desires only render activities rational or

\(^{65}\) Ibid. p. 160.
irrational when considered in tandem means that adjustments to either can be balanced by adjustments to the other in order to maintain a high degree of rationality across competing interpretations. Davidson deems the resulting number of alternative explanations for a subject’s activities “unacceptably large”, so he demands a further means for constraining interpretations.

Here, Davidson appeals to Ramseyan decision theory. He claims that, under idealized conditions, the theory of choice under uncertainty allows for unique specification of thoughts. But in order to apply decision theory, we need some way of judging our subject’s preferences, some way of disentangling the roles played by desires and beliefs in the determination of actions. And according to Davidson, “all the standard ways of testing theories of decision or preference under uncertainty rely on the use of language.”

Davidson’s worry here has to do with one of the seven requirements for thought attribution that I identified above: the intensionality requirement. It might be easy enough to determine which of two objects a non-linguistic subject prefers; we could allow the subject access to either of the objects, but not both, and see which object the subject selects. But this will not give us information about which specific aspects of the selected object made it preferable: “A man who has chosen an apple rather than a pear when offered both may be expressing a preference for what is on his left rather than his right, what is red rather than yellow, what is seen first, or judged more expensive.” In order to assess the subject rationally, we must figure out not just which objects the

---

66 Ibid. p. 160.
67 Ibid. p. 163.
subject is thinking about, but also the categories under which the subject conceives of the object as it makes its selections. Otherwise, we will still be left with the wide variety of belief-desire combinations that the holism of the mental allows.

So, Davidson’s point here is that practically, we cannot work out all of the holistic issues regarding belief and desire even in the presence of language. Under idealized conditions, an interpreter could apply Ramseyan methods\textsuperscript{68} to solve this problem. But they require us to make distinctions between a subject’s preferences at the fine-grained level of properties rather than the coarse-grained level of objects, and all of our standard ways of making such distinctions rely on linguistic behaviors by the subject. Language allows us to structure the choice situation in order not just to see what the subject has chosen, but to understand which features guided the subject’s choice.

So, we can draw the following argument out of Davidson’s considerations.

\textit{Davidsonian Holism Argument:}

1 If Ramseyan methods cannot be applied, then the holism of the mental leaves open an unacceptably large number of competing interpretations.
2 If we cannot judge subjects’ selections at the grain of intensions, then we cannot apply Ramseyan methods.
3 If a subject is non-linguistic, then we cannot judge the subject’s selections at the grain of intensions.
4 So, if a subject is non-linguistic, then the holism of the mental leaves open an unacceptably large number of competing interpretations.\textsuperscript{69}

Davidson holds that his appeal to Ramseyan methods strongly suggests that thought attributions and linguistic interpretation must go together, but that “it remains to

\textsuperscript{68} One gets the impression that Davidson believed that decision theory codified the methods of judgment upon which interpreters had been relying implicitly. See Davidson (1974a, 1985).
\textsuperscript{69} For more information on the structure of this derivation and the others in this chapter, see the Appendix to this dissertation.
say … why the attribution of thought depends on the interpretation of speech.” He continues: “The general … reason is that without speech we cannot make the fine distinctions between thoughts that are essential to the explanations we can sometimes confidently supply.” His point here also concerns intensionality. He points out that our everyday attributive practice cares about intensionality, as it must if it is to allow us to judge beings by rational standards. A person who believes that Scott is not the author of *Waverley* might be perfectly rational, whereas a person who believes that Scott is not Scott has failed to properly apply inferential standards concerning identity.

Thus, when we attribute thoughts to our linguistic peers, we strive to capture not only the objects of their thoughts, but also the ways in which they think about those objects. With non-linguistic beings, we face a challenge: “The dog, we say, knows that his master is home. But does it know that Mr. Smith (who is his master), or the president of the bank (who is that same master), is home? We have no real idea how to settle, or make sense of, these questions.” Whereas the rich behavior provided by language use enables attribution of the dense, extensive network of beliefs that would allow us to make sense of and settle these questions, non-linguistic behavior might be too sparse.

In practice, language often makes the interpretive task even easier. When the subject speaks our language, we can simply model the attribution after a sentence the subject would accept. When the subject speaks a different language, we can model the attribution after a sentence that translates a sentence the subject would accept. But when

---

70 Ibid. p. 163.
71 Ibid. p. 163.
72 Ibid. p. 163.
the subject speaks no language, what are we to do? It seems clear to me that we should not say the dog thinks the president of the bank is home, as we can safely assume that the dog has no concept of banks, or presidents, for that matter; I will support this intuition with a principled way to rule out such attributions in chapter 4. But even if we know we don’t want to describe Mr. Smith as the president of the bank for the purpose of attributing thoughts about him to the dog, how should we describe him for that purpose?

We might try to model the attribution after a sentence that the dog would accept, were the dog to speak our language. Maybe, but that suggestion contains a counterfactual that at best would have to be spelled out carefully, and at worst would involve appeal to possible worlds where the subject’s cognitive processes are not even comparable to the actual processes we are trying to interpret. We would then have an account of what the subject would think if it had language, rather than an account of what it does think.

Similar difficulties beset attempts to find other structural features in the cognition of non-linguistic subjects. Davidson mentions the distinction between universals and conjunctions, conditional thoughts, and thoughts with mixed quantification as three examples. I will grant that at least the first two pose difficulties and warrant a response. (I am unsure why we would feel pressure to make sense of mixed quantification for non-linguistic subjects.) I will discuss universals and conditionals in further detail later in this chapter.

The considerations about specificity of attributions seem to constitute another argument against the possibility of non-linguistic thinkers.

---

73 Ibid. p. 164.
Davidsonian Network Argument:

1 The specific conceptual composition of a thought is determined by its place in a dense, extensive network of beliefs.
2 Only linguistic subjects behave in ways rich enough to embody a dense, extensive network of beliefs.
3 So, only linguistic subjects can have thoughts with determinate conceptual contents.

Davidson himself, though, does not think that these considerations constitute an argument against the possibility of non-linguistic thinkers. Impressive as he finds the considerations he has discussed, he nonetheless grants that an opponent might view this as either a merely epistemic problem or a mere matter of our discomfort with departure from the familiar features of linguistic interpretation. Thus, he writes of interpretation of non-linguistic subjects, “the evidence will not be adequate to justify the fine distinctions we are used to making in the attribution of thoughts … our attributions will be seriously underdetermined in that many alternative systems of attribution, many alternative explanations, will be equally justified by the available data.”74 So, he only complains that we are “used to making” fine-grained attributions, and not that we must make them, and that our justification for choosing one interpretation rather than many others will be lacking, and not that there is no correct interpretation. He is hedging on his earlier claim that the number of interpretations left open due to the holism of the mental must be considered unacceptably large. He seems to find it unacceptably large himself, but he

---

74 Ibid. p. 164.
also refrains from relying on that point, apparently thinking that he will not be able to
defend it to an opponent’s satisfaction.

Perhaps he hedges here because of his own appeal to Ramsey’s methods, and the
fact that even those methods require idealization in order to eliminate excessive
underdetermination. If we must idealize anyway even in the case of linguistic beings,
then perhaps there is room for an opponent to hold that we must merely idealize a little
more to make sense of non-linguistic beings. Whatever the reason for the hedging,
Davidson seeks a better argument.

He briefly considers a line from *Word and Object*,75 in which Quine maintains
that the act of attributing an attitude is an act of mimicry. The attributer imitates an
actual or possible speech act of the subject. But while this is a picturesque way of
presenting the issue, it does not advance the case against non-linguistic thinkers beyond
the aforementioned difficulty with choosing an acceptable attribution for a non-linguistic
subject. We are once again left to wonder how far we can stretch the notion of a
“possible speech act” of a being that is non-linguistic in the actual world. Furthermore,
Davidson grants that an opponent might simply reject the Quinean account of attitude-
attributing actions.

At this point, Davidson shifts to a different kind of argument, one based on
objectivity and the source of the contrast between objective truth and subjective belief.
He argues that the concepts of objective truth and error emerge only in the context of
linguistic interpretation, that one must grasp the concepts of objective truth and error if

---

one is to have the concept of belief, and that only one who has the concept of belief can have thoughts. This line appeals to the subject’s status as a linguistic interpreter, so it promises to defend the claim that Davidson called his chief thesis: that only interpreters of speech can be thinkers.

Davidson begins his argument by pointing out that “a sentence is held true because of two factors: what the holder takes the sentence to mean, and what he believes. In order to sort things out, what is needed is a method for holding one factor steady while the other is studied.”76 With a linguistic subject, we can determine which sentences the subject holds true under various circumstances. We can then use this information to work out a theory of the held-true for the subject that meets Tarskian formal constraints on truth. Davidson has argued at length that a theory of truth for a language that meets Tarski’s constraints can serve as a theory of meaning for that language’s expressions.77 So, according to Davidson, a theory of the held-true for a subject’s language can serve as a theory of what the subject’s expressions mean, insofar as the held-true corresponds to the true. With such a theory, one can in turn sharpen the tests one uses to determine what the subject holds true. That will improve the theory of meaning, and so our understanding of the subject will continue to be honed.

But that method works only insofar as the held true corresponds to the true, and “of course it cannot be assumed that speakers never have false beliefs. Error is what gives belief its point.”78 Nonetheless, Davidson argues that we can take for granted that

77 See especially Davidson (1967).
most of the subject’s beliefs are correct. Here, Davidson makes an important point. We identify beliefs by the place they occupy in a pattern of beliefs. For example, part of what allows us to identify a belief that snow is white and grass is green is the inferential relationship it bears to the beliefs that snow is white and that grass is green. These in turn are partially identified by their relationships to other beliefs about snow, grass, white things, and green things. An isolated error will stand out against a dense background of true beliefs, because there will still be ample connections between the true beliefs to allow us to place them in the network, and discern their identities. In the absence of such a background, our ability to even assign content to a belief is jeopardized. Rather than attributing massive error to the subject, we will simply find that we cannot successfully interpret the subject at all. So, if we find that we are able to interpret a subject at all, then we can safely conclude that the subject has mostly true beliefs.

Of course, the interpreter must rely on her own ideas about what is true in order to judge the subject, and these may also be false on occasion. So, the attainable goal of the interpreter is to optimize agreement between one’s subject and oneself. To succeed as an interpreter, then, one must have some grasp of the distinction between truth and error, and an understanding that disagreements between oneself and the subject will often indicate that at least one party is in error.

On the basis of these considerations, Davidson concludes that, “we should acknowledge that the concepts of objective truth, and of error, necessarily emerge in the

---

79 Or, more precisely, to minimize inexplicable disagreement.
context of [linguistic] interpretation.\textsuperscript{80} Note the crucial modal ambiguity in Davidson’s claim. On one disambiguation, the claim is that the concepts of objective truth and error are incapable of emerging outside of the context of linguistic interpretation. A subject’s grasp of those concepts would then depend on the subject’s ability to occupy the role of linguistic interpreter, accomplishing one inferential step toward defense of Davidson’s chief thesis.

But the discussion that follows seems to indicate a different disambiguation: that the context of linguistic interpretation forces the emergence of the concepts of objective truth and error. Davidson continues, “The distinction between a sentence being held true and being in fact true is essential to the existence of an interpersonal system of [linguistic] communication, and when in individual cases there is a difference, it must be counted as error.”\textsuperscript{81} That is correct, as far as it goes. Disagreements between interpreter and interpretee are bound to occur. When a disagreement occurs within a context governed by an interpersonal, shared standard such as truth, we are forced to count at least one side of the disagreement as an error. Both sides cannot meet the standard in such cases, so at least one must be wrong. But the arrow of implication points the wrong way for Davidson’s argument. These considerations indicate: “If a subject is a linguistic interpreter, then the subject grasps the concepts of truth and error.” Davidson needs: “If a subject grasps the concepts of truth and error, then the subject is linguistic.” So, more work is required.

\textsuperscript{80} Ibid. pp. 169-170.
\textsuperscript{81} Ibid. p. 170. I have inserted the word “linguistic” into the quote because, in context, it is clear that Davidson intends to refer only to linguistic systems of communication.
Here, Davidson appeals to the role played by belief in the method of linguistic interpretation. He holds that the notion of belief is given shape by its relationship to meaning, truth, and the held-true within that method. His argument is less than explicit, but it goes something like this. The process of interpretation forces upon us the notions of truth and error, and thus a distinction between expressions that are true and expressions that are merely held true. We need to make sense of many discrepancies as genuine disagreements rather than just episodes of talking past each other, because we undermine the intelligibility of our interpretation of a subject if we posit too much incommensurability between their views and our own. The meanings of expressions must remain stable in order to make sense of disagreement as such, rather than just as talking past each other. Since we cannot appeal to differences in meaning in order to make sense of genuine disagreements, the difference between the held true and the true cannot be explained in terms of differences in meaning. So, there must be some other factor that explains the difference: a subjective state that partially determines which sentences are held true, but is judged according to an independent standard of truth. Davidson sums up: “The concept of belief thus stands ready to take up the slack between objective truth and the held true...”

That is an interesting account of why the notion of belief must play an active role in the context of interpretation. The methodology of interpretation forces us to posit something that plays a specific role---the role of belief---in order to make sense of meaning, the held-true, and truth. So linguistic interpreters must have the notion of a

---

82 Ibid. p. 170.
belief. But the arrow of implication still points in the wrong direction for Davidson’s purposes.

Davidson follows up “The concept of belief thus stands ready to take up the slack between objective truth and the held true…” with “…and we come to understand it just in this connection.” He claims that belief, as a private attitude, is only intelligible “as an adjustment to the public norm provided by language.”

So, he is attempting to get the needed direction of implication by holding that only linguistic practice provides a normative context in which we can make sense of the role of belief. So, the idea seems to be the following.

*Davidsonian Error Argument Schema:*

1. Belief is only made intelligible in contexts involving error.
2. The contrast between error and correct conduct exists, or is made intelligible, only in a norm-governed situation (in the case of belief, the relevant norm is set by the standard of objective truth).
3. Objective truth exists, or is made intelligible, only in contrast to the held true.
4. The notion of being held true applies only to sentences, or is made intelligible only in the context of its application to an interpretee’s sentences.
5. Only beings to whom belief is intelligible can have a belief.
6. A being must have beliefs in order to have any thoughts.
7. So, only linguistic interpreters can have thoughts.

Note that, due to the disjunctive nature of lines 2, 3, and 4, we have potentially eight different versions of this argument. I will leave assessment of these lines for Chapter 4. Any of them, if successful, would render a subject’s ability to grasp the notion of belief dependent on the subject having interpreted the sentences of others.

---

83 Ibid. p. 170.
Given that beliefs must be attributed in order to make sense of attributions of any thoughts, the sole remaining step for Davidson is to argue that only a being with the concept of belief can have a belief. That would suffice to show that only an interpreter of language could be a thinker.

To this end, Davidson writes: “Can a creature have a belief if it does not have the concept of a belief? It seems to me it cannot, and for this reason. Someone cannot have a belief unless he understands the possibility of being mistaken, and this requires grasping the contrast between truth and error---true belief and false belief.”\textsuperscript{84} So, having a belief depends on understanding that one could be mistaken, and this depends on grasping the contrast between truth and error.

This way of finishing the argument renders the structure of Davidson’s argument hazier. Instead of a sub-argument that having a belief depends specifically on having the concept of a belief, we get instead a sub-argument that having a belief depends on having the network of associated concepts that Davidson maintains arise only in the context of linguistic interpretation. Perhaps that is all we really need to know, then: Davidson thinks that a cluster of concepts---among them belief, truth, and error---are only grasped by linguistic interpreters due to their interlocking roles in the process of linguistic interpretation.\textsuperscript{85}

Note that Davidson’s key claim---that having a belief depends on understanding the possibility of being mistaken---goes undefended in “Thought and Talk”. But I suspect that some version of the reflection requirement drives Davidson’s intuitions. The

\textsuperscript{84} Ibid. p. 170.
idea would be that beliefs are, by nature, the kinds of states that can combine with desires in reflective decision making. So, only beings that can reflectively deliberate can have states that function as beliefs. And part of reflective deliberation involves taking one’s beliefs for what they are: subjective states that do not always meet the standard of truth. If so, we would have something like this.

*Davidsonian Concept Cluster Argument:*

1. The concepts BELIEF, TRUTH, and ERROR are acquired only as a group.
2. That group of concepts is acquired only in the process of linguistic interpretation.
3. Grasp of the concepts BELIEF, TRUTH, and ERROR is required in order for a being to be in states that function as beliefs.
4. A being must have beliefs in order to have any thoughts.
5. So, only linguistic interpreters can have thoughts.

My suspicion that Davidson reasons along these lines gains support from his elaboration on these considerations in “Rational Animals”. In that paper, Davidson discusses another phenomenon that he takes to require reflective awareness: the phenomenon of surprise. He also introduces a key feature of his later thought on these topics, an analogy to geometric triangulation. He employs the triangulation metaphor to help elucidate an aspect of his philosophy that is introduced only at the very end of “Thought and Talk”: the idea that the contrasting notions of subjective and objective take shape alongside the notion of the intersubjective. Additionally, Davidson elaborates on the idea that thoughts can be identified only against the backdrop of a rich network of beliefs. These three ideas—the need for a rich network of beliefs, the role of triangulation in explaining the intersubjective, the objective, and the subjective, and the
reflective nature of surprise as it relates to belief---provide the impetus for Davidsonian arguments in “Rational Animals” that go beyond those of “Thought and Talk”.

Considerations that Davidson Marshals in “Rational Animals”

In “Rational Animals”, Davidson rehearses many themes familiar from “Thought and Talk”. We are reminded about the holistic nature of the attitudes, about the difficulty in specifying the manner in which a dog might think of objects, and that one can only identify a thought by placing it in a network of beliefs. But we are given more clues as to which aspects of the network Davidson thinks will be absent for non-linguistic beings. In “Thought and Talk”, Davidson demanded that there be “endless interlocked beliefs” in order to support identification of a thought. The demand that the beliefs be endless could be a demand that the structure of thought mirror the structure of language, which allows for endless generation of new sentences from finitely many words. The mental analogs of words, concepts, would have to allow for endless generation of new thoughts in order for thought to be parallel to language. One gets the impression that the density and extent of the network of beliefs are the important factors for thought identification, and that Davidson’s complaint is that dogs do not behave in rich enough ways to allow us to ascertain whether they could be credited with networks that are vast enough and dense enough to support attributions of specific concepts.

---

86 See p. 157.
But in “Rational Animals”, Davidson indicates that the kinds of beliefs that comprise the network are important, too. Davidson asks whether a dog can believe of an object that it is a tree, and answers that it could do so only if it had many general beliefs about trees. So, it is not enough that the dog have many beliefs about particular trees. Even if the dog’s beliefs about particular trees could be expanded endlessly, the dog would still need to have beliefs about trees in general in order for Davidson to credit it with any beliefs that contain the concept TREE. The dog must recognize the category of trees, over and above recognizing many individual members of the category.

If the dog had a word for trees, we would be in position to attribute to the dog general beliefs about trees. Given that the dog does not have a word for trees, how are we to tell whether it recognizes the category? Here, one of Davidson’s points from “Thought and Talk”87 takes on new life: the behavior of non-linguistic subjects gives us no good way to distinguish between universal thoughts and conjunctive thoughts. Truly general beliefs about trees would have to be universal in character; a conjunction of beliefs about particular trees could be accomplished without any recognition of the category of trees. So, we have a distinct line of argument here.

Davidsonian Generality Argument:

1 The specific conceptual composition of a thought is determined by its place in a network of beliefs.
2 The network must contain general beliefs if it is to identify the specific conceptual composition of a thought.
3 Only when a subject uses categorical terms do we have any way to make sense of the distinction between general beliefs and conjunctions of beliefs about particulars.

87 See p. 164.
4 So, only linguistic subjects (i.e., subjects who use categorical terms) can have thoughts with determinate conceptual contents.

As with the Network Argument, Davidson backs down from endorsing the Generality Argument as sufficient to rule out the possibility of non-linguistic thinkers. He claims instead that this “only shows that there can’t be much thought without language.”

That cannot be what Davidson really means. He has just argued at length that in order for there to be any thought, there must be a vast network of thoughts. So, it makes no sense for him to claim that he has only shown that there cannot be much thought without language. Rather, he seems to be acknowledging that his opponents, by rejecting or softening one of his premises, might hold that he has only shown that there cannot be much thought without language.

Whatever the reason, the line is abandoned in favor of the official master argument of “Rational Animals”: “The argument has two steps. … First, I argue that in order to have a belief, it is necessary to have the concept of belief. Second, I argue that in order to have the concept of belief one must have language.” As Brandom points out, this is really more an argument schema than an argument. And indeed the remainder of “Rational Animals” is spent filling in the motivations for the two premises in ways that suggest sub-arguments. So I will offer several more complex arguments that contain Davidson’s two official steps as sub-conclusions.

---

89 Ibid. p. 478.
Arguments for sub-conclusion 1

I will first consider Davidson’s arguments for sub-conclusion 1, the claim that in order to have a belief, it is necessary to have the concept of a belief. It is in this context that the phenomenon of surprise becomes important. Davidson discusses an exchange between Norman Malcolm and Donald Weiss. Malcolm distinguishes believing that p from having the thought that p, holding that the former requires only awareness that p, whereas the latter requires awareness of awareness that p.91 This is somewhat problematic, since “is aware that” seems to have both factive and phenomenological aspects that “believes that” lacks. But the point is clear enough: Malcolm maintains that non-linguistic beings can have a non-reflective capacity to believe that p, but not a reflective version of that capacity.

Davidson’s goal at this point in “Rational Animals” is to argue that believing that p depends on being able to conceive of one’s beliefs as such. That argument is supposed to pair with one that denies the latter capacity to non-linguistic beings. But Davidson does not offer a direct argument. Instead, he discusses Weiss’s proposed example of a non-linguistic thinker. We are to imagine a being, Arthur, who has hatched and grown to adulthood in isolation from other beings, guaranteeing that he has not taken part in linguistic interactions. Apparently, Arthur’s abode provides him with all that is essential to living, and so he never ventures beyond it into a world in which he could interact with

91 Malcolm (1972-3).
other beings. The dwelling also (conveniently) has walls that are reflective when viewed from the inside, but transparent when viewed from the outside, so we are able to watch Arthur’s every move without his being any the wiser.

Weiss describes a scenario that is supposed to convince us that we should attribute thought to Arthur. Over some time, we have observed Arthur placing various metals in his fire, then hammering them into different shapes. One day, Arthur puts a new metal in the fire, but his hammering does not change the metal’s shape. Arthur behaves in ways typical of a person confronted with a puzzle. He seems agitated, he paces, he sits with eyes fixed straight ahead, etc. Finally, he leaps up, piles more wood on his fire, plunges the metal back in, and hammers it into a different shape. No longer looking agitated, he goes about his daily activities.

Weiss thinks that Arthur’s observers are now justified in attributing thought to Arthur. Arthur apparently believed a generalization such as, “All metals become malleable when heated in the fire.” We observers are to glean from Arthur’s posture and conduct that he has become befuddled upon encountering an instance that does not conform to the generalization. We are also supposed to conclude that Arthur mulled over the problem and leapt up when the solution, a hotter fire, occurred to him.

I will give my own assessment of Arthur in Chapter 4. For now, note that Davidson’s view does not permit him to attribute thought to Arthur. No matter how complex or suggestive of thought Arthur’s behaviors might be, Arthur is non-linguistic, so he should not be capable of thought, according to Davidson. Davidson has two

---

92 Weiss (1975).
options with respect to Arthur. He can either hold that Arthur’s behaviors do not warrant attribution of thought, or he can hold that the behaviors Arthur is supposed to display could only be displayed by a linguistic being---i.e. no being like Arthur could really exist.

But Davidson does not pass either judgment on Arthur. Instead, he focuses on one aspect of Weiss’s description of the scenario. Arthur appears to be surprised when the new metal fails to become malleable. According to Davidson, “it is essential [for attributing thought to Arthur] that we be able to describe Arthur as being surprised … what I think is clear is that if he is surprised, he does have reflective thoughts…”93 Note that “reflective” here does not just mean that the thought process was cautious and time-consuming, rather than careless and spontaneous. To be reflective in Davidson’s sense, a thought process must be at least partially about one’s thoughts. This requires that one be able to conceive of one’s own thoughts as such. Recall that Davidson’s goal at this point is to defend the claim that having beliefs depends on having the concept of belief. So, when Davidson claims that Arthur’s being surprised entails that Arthur has reflective thoughts, we can understand Davidson to be looking for an argument that fits the following schema:

(Premise 1) To be subject to surprise, one must have the concept of belief.
(Premise 2) To be a thinker, one must be subject to surprise.
(Conclusion) To be a thinker, one must have the concept of belief.

Premise 1 is a rephrase of Davidson’s claim that Arthur’s being surprised entails that Arthur has reflective thoughts. The conclusion is premise one of Davidson’s official

---

93 Davidson (1986c), p. 479.
master argument schema from “Rational Animals”. I supply Premise 2 because it is the quickest way to get from Premise 1 to the conclusion.

But the premises are not obvious and must first be defended as subconclusions. Davidson begins by pointing out that surprise, in contrast to merely being startled, requires beliefs. That is reasonable: one has to believe the world is one way in order to be surprised to find that it is some other way. Davidson illustrates his point with an example in which he believes there is a coin in his pocket, but is surprised upon emptying his pockets to find that there were no coins. He draws the lesson that surprise requires belief, and continues, “And perhaps it is equally clear that having a belief, at least one of the kind I have taken for my example, entails the possibility of surprise.”

A belief that there is a coin in one’s pocket falls under many kinds. Davidson does not say which of these he has in mind. Some beliefs, such as the belief that $2 + 2 = 4$, by their nature prevent us from ever finding out that they are false, let alone ever being surprised to find that they are false. Perhaps the hedge away from full generality is meant to guard against counterexamples that exploit certain beliefs’ inherent inimicality to revision in order to show that there are beliefs about which it need not be possible to be subject to surprise at the need for revision.

Davidson seems to support this interpretation when he follows up the sentence in question with: “If I believe I have a coin in my pocket, something might happen that would change my mind.” If my interpretation is right, then the kind of belief Davidson has in mind is something like “beliefs that one could readily discover to be false”. And

---

94 Ibid. p. 479.
95 Ibid. p. 479.
Davidson’s defense of premise 2---to be a thinker, one must be subject to surprise---would be that beliefs that one could readily discover to be false entail the possibility of surprise, and that one cannot be a thinker without having some beliefs of that kind.

Davidson’s defense of premise 1---to be subject to surprise, one must have the concept of belief---relies on the fact that surprise does not merely involve alteration of beliefs. Davidson writes: “Surprise requires that I be aware of a contrast between what I did believe and what I come to believe. Such awareness, however, is a belief about a belief: if I am surprised, then among other things I come to believe my original belief was false.”

So, Davidson claims that surprise requires an awareness of a contrast between something that was believed and something one comes to believe, and that this must involve reflection on one’s beliefs as such.

This is enough to complete an argument that thinkers must have the concept of belief.

Davidsonian Master Argument Subargument 1, version 1:

1 To have a belief that one could readily discover to be false entails that one is potentially subject to surprise.
2 One cannot be a thinker without having some beliefs that one could readily discover to be false.
3 Surprise requires an awareness of a contrast between something that was believed and something one has come to believe.
4 Awareness of a contrast between something that was believed and something one has come to believe requires grasp of the concept of belief.
5 So, to be a thinker, one must have the concept of belief.

---

96 Ibid. p. 479.
97 I call this “version 1” for reasons that will become clear presently.
But Davidson complicates matters when he writes: “I do not need to insist that every case of surprise involves a belief that my original belief was false (though I am inclined to think so).”\textsuperscript{98} So, Davidson seems to suggest he can weaken claim 3 and still push the inference through. He continues: “What I do want to claim is that one cannot have a general stock of beliefs at all without being subject to surprises that involve beliefs about the correctness of one’s own beliefs.”\textsuperscript{99} So, Davidson is telling us that surprise does not always involve reflection, as line 3 claims, but he is also giving us the new information that a being cannot have a general stock of beliefs unless it is sometimes subject to the reflective kind of surprise. So, one would want to offer the following argument, with 3 appropriately altered and the new information concerning the relationship between reflective surprise and a general stock of beliefs introduced at line 5:

1 To have a belief that one could readily discover to be false entails that one is potentially subject to surprise.
2 One cannot be a thinker without having some beliefs that one could readily discover to be false.
3 Reflective surprise requires an awareness of a contrast between something that was believed and something one has come to believe.
4 Awareness of a contrast between something that was believed and something one has come to believe requires grasp of the concept of belief.
5 To have a general stock of beliefs, one must be subject to reflective surprise.
6 To be a thinker, one must have the concept of belief.

But that argument is invalid; 6 does not follow from any combination of the previous lines. One might try to patch it up by appealing to Davidson’s well-defended claim that

\textsuperscript{98} Ibid. p. 479.
\textsuperscript{99} Ibid. p. 479.
to be a thinker requires a general stock of beliefs. That claim, along with lines 3-5, does yield 6 (I have kept the initial numbering for ease of reference):

Davidsonian Master Argument Subargument 1, version 2:

3 Reflective surprise requires an awareness of a contrast between something that was believed and something one has come to believe.
4 Awareness of a contrast between something that was believed and something one has come to believe requires grasp of the concept of belief.
5 To have a general stock of beliefs, one must be subject to reflective surprise.
* To be a thinker, one must have a general stock of beliefs.
6 To be a thinker, one must have the concept of belief.

Unfortunately, lines 1-2 from version 1 play no role at all in version 2. And much of Davidson’s discussion was spent motivating lines 1-2; the premise at line * has not been motivated at all by Davidson’s discussion. So, although I will consider version 2 in Chapter 4\footnote{The versions in Chapter 4 and the appendix are revised to eliminate the awkward wording and numbering that I use here for the sake of exposition.}, my suspicion is that Davidson was mistaken when he thought he could get by with a weakened version of the original line 3. The inferential pattern he had been building toward is ruined when 3 is weakened. One would hope that further information would save Davidson from this problem, but at this point, he declares this phase of his argument complete.\footnote{Ibid. p. 479.}

I believe that the example of surprise actually obscures the point Davidson wants to make. Davidson wants to establish that a non-reflective capacity to believe that p depends on a reflective capacity to consider one’s beliefs as such. He identifies a
phenomenon that he believes is reflective, surprise, and tries to show that having beliefs entails that one may be subject to the purportedly reflective phenomenon. But he does not wish to hold that thought must always be reflective, only that the capacity for non-reflective thought depends on the capacity for reflective thought. This, I think, leads to his mistaken claim that he need not insist that surprise is always reflective in order for his argument to work. But if surprise is not always reflective, then an argument that belief entails the possibility of surprise will not do the work Davidson wants to be done.

If we drop the talk of surprise, and take Davidson to be claiming that a being cannot have a general stock of beliefs without being reflective, a new possibility for defending the target claim emerges. Can a being really have states that function as beliefs if the being does not ever reflect on any of those states? The idea would be that any given belief need not be reflected upon to play its role, but that our ability to recognize a creature as non-reflectively employing a belief depends on our ability to recognize that creature as reflectively employing some beliefs. If a being is unable to reflect on its own cognition, it is not able to take control of its actions. If a being cannot exercise deliberate control over its actions, then it does not have cognitive states that function as beliefs do, for beliefs are partially defined by the role they play in deliberate action.

The point of Davidson’s talk of surprise was that, according to Davidson, a being that is subject to surprise is capable of comparing its own beliefs. That means that the being can focus its attention on its beliefs. And that is an essential component of the ability to consider how those beliefs might advise one to act in order to accomplish one’s
goals. And thinkers, one might hold, must be able to do that. So, we have another
potential argument for the claim that thinkers must have the concept of belief:

*Davidsonian Master Argument Subargument 1, version 3:*

1. Thinkers must be able to take deliberate control of their actions.
2. In order to take deliberate control of their actions, beings must consider
   how their beliefs advise them to act in the pursuit of their goals.
3. In order to consider how their beliefs advise them, beings must be able
   to focus their attention on their beliefs.
4. In order to focus attention on one’s beliefs, one must be able to conceive
   of one’s beliefs as such.
5. Thus, thinkers must be able to conceive of their beliefs as such.

In this case, I cannot claim that I am simply explicating arguments that Davidson
has given. His discussion does not focus on these issues. But thoughts like these would
make more sense of his discussion of surprise. A concern for the manner in which the
role of belief determines the nature of belief is shown by Davidson’s contentions that
beliefs typically expose believers to the possibility of surprise. Because Davidson thinks
that surprise is reflective, one could see him as attempting to connect the nature of belief
to the role it plays in reflection. He also emphasizes the connection between belief and
the role it plays in the process of practical reasoning and intentional action.\footnote{Ibid. p. 476.} Since he
takes participation in that process to demand some capacity for reflection, that would
explain why he attempted to finish the argument by making an appeal to specifically
reflective surprise as a condition on having beliefs generally. So, version 3 might not be
an argument Davidson ever gave, but it fits within a Davidsonian view of thought, and
arguably explains the motivations behind the versions Davidson did offer.
Having defended the claim that thinkers must have the concept of belief, Davidson turns his attention to the claim that only language users have the concept of belief. In “Thought and Talk”, Davidson discussed the contrast between something subjective, the held-true or believed, and something objective, the truth. In “Rational Animals”, he attempts to connect the dawning of these notions to something intersubjective, the shared standards of linguistic interpretive practice.

Davidson reminds us of a claim he defended in “Thought and Talk”: that a key part of the role of beliefs is that they are subjective states that are held to an objective standard, truth. The practical role played by the notion of belief is that it is the concept of a state that occupies one of the two roles needed in order to make sense of a subjective-objective dichotomy; beliefs are the subjective states that provide a contrast with objective truth. So, sense is made of the contrast between subjective and objective by opposing subjective belief to objective truth, and the concepts of belief and truth take shape only through the roles they play in making sense of the subjective-objective contrast. Grasp of the subjective-objective contrast implies grasp of the dual concepts of belief and objective truth, and vice versa.

Next, Davidson claims that linguistic communication suffices to demonstrate grasp of the subjective-objective contrast. Clearly, he wishes to connect the origins of the

---

103 Ibid. p. 479.
subjective-objective contrast with the need to treat oneself as operating within a shared world, accountable to an intersubjective standard of truth. Linguistic communication involves entertaining the same propositions, taking each other to be doing so, and judging whether the world satisfies those propositions’ truth conditions or not. The need to treat multiple subjects as moving in the same world imposes a shared standard of truth. That intersubjective standard demands that propositions be measured against conditions that obtain objectively in the shared world. So, the intersubjective world must also be taken to be an objective world. Grasp of the intersubjective standard of truth used in linguistic communication thus implies grasp of the contrast between subjectively entertained propositions and objective truths.

So, Davidson has motivated several claims. He has argued that only beings that grasp the contrast between subjective and objective can have the concept of belief. He has pointed out that linguistic communication requires interlocutors to share a standard of truth, and to treat themselves as responding to an intersubjective world. So, linguistic beings must be credited with the concepts of intersubjective truth and intersubjective world. And he has called attention to the ways in which application of a shared standard of truth within linguistic practice demands grasp of the idea that propositions are true or false on the basis of conditions that obtain objectively within the shared world. So, grasp of the notion of intersubjective truth entails grasp of the subjective-objective contrast.

So, we have an argument like the following:

Preliminary version of Davidsonian Subargument 2:

1 When an intersubjective standard of truth is applied to propositions about a shared world, the propositions are recognized to be
subjectively entertained, and judged true or false depending on conditions that are recognized to obtain objectively within the shared world.

2 Linguistic communication involves applying an intersubjective standard of truth to propositions about a shared world.

3 Grasp of the contrast between subjective and objective is embodied in the opposition of the dual concepts of subjective belief and objective truth.

4 So, linguistic communication entails grasp of the concept of belief.

But what Davidson really wants to be able to claim is not just that language use guarantees grasp of the concept of belief (that could have been demonstrated more straightforwardly by pointing out that many utterances express beliefs), but that only language users can have the concept of belief. That is the target claim needed to finish the argument that only language users can be thinkers. And our preliminary version of Davidsonian Subargument 2 falls short of establishing the target claim in two ways. The first problem occurs at line 1. It is not enough that application of the concept of intersubjective truth guarantees grasp of the subjective-objective contrast. Instead, Davidson wants to argue that only through grasp of the concept of intersubjective truth could one arrive at grasp of the subjective-objective contrast. The second problem occurs at line 2. It is not enough that linguistic communication requires application of a standard of intersubjective truth. Davidson needs grasp of the concept of intersubjective truth to depend on linguistic practice. So what Davidson really wants to argue is:

**Davidsonian Master Argument Subargument 2, version 1:**

1 Only through grasp of the concept of intersubjective truth could one come to grasp the subjective-objective contrast.
2 Grasp of the concept of intersubjective truth depends on participation in linguistic communication.
3 A being can make sense of the dual concepts of subjective belief and objective truth only in their roles as embodiments of the contrast between subjective and objective.
4 So, only those who have communicated linguistically can grasp the concept of belief.

Of the three new premises, only line 3 has been defended at this point in “Rational Animals”. With respect to line 2, Davidson offers only the following: “And perhaps it is plausible enough that having the concept of intersubjective truth depends on communication in the full linguistic sense.”¹⁰⁴ Perhaps this is supposed to be plausible because propositions put forward in linguistic interaction would be the items to which one would first come to apply a standard of intersubjective truth. As Davidson gives no further explanation, we are left to imagine our own reasons why this claim might be regarded as plausible enough.

As for line 1, Davidson confesses that he does not know how to establish that only grasp of an intersubjective standard of truth could provide a being with grasp of the subjective-objective contrast. But he also states that he has no idea how else one might arrive at the notion of an objective truth. Here, he appeals to the triangulation metaphor. First, he gives an example involving geometrical triangulation:

If I were bolted to the earth I would have no way of determining the distance from me of many objects. I would only know they were on some line drawn from me toward them. I might interact successfully with objects, but I could have no way of giving content to the question of where they were. Not being bolted down, I am free to triangulate.¹⁰⁵

¹⁰⁴ Ibid. p. 480.
¹⁰⁵ Ibid. p. 480.
Triangulation is a process in which the distance to a point is determined by comparison to two other accessible points. If one connects the points into a triangle, one can use the measure of the distance between the two accessible points and the measure of the angles that branch from the two accessible points to calculate the distance to the third point. If Davidson were bolted to the earth, he would have access only to one point, and not a second. So, he would be unable to use the triangulation method to determine the distance between himself and other items.

Having explained triangulation, Davidson then draws an analogy between the use of triangulation to determine distances and the role played by multiple perspectives within linguistic interaction in the development of the notion of objectivity:

Our sense of objectivity is the consequence of another sort of triangulation, one that requires two creatures. Each interacts with an object, but what gives each the concept of the way things are objectively is the base line formed between the creatures by language. The fact that they share a concept of truth alone makes sense of the claim that they have beliefs, that they are able to assign objects a place in a public world. 106

So, Davidson has offered several features of linguistic interaction that parallel features of geometrical triangulation. The two creatures in the linguistic scenario are like the two accessible points in the geometrical scenario. The object is the third point. Each creature’s interactions with the object and each other are like the degrees of the angles. The intersubjective standard of truth is supposed to be like the measure of the distance between the two accessible points. Knowledge of the measures of angles, along with the

106 Ibid. p. 480.
measure of the base distance, suffice to place the third point; the two creatures’
interactions with each other and the object, along with the intersubjective standard of
truth that is operative because the interactions are linguistic, suffice to identify an object
of thought in an objective reality.

A lone creature, on the other hand, would be comparable to Davidson being
bolted down. Bolted-down Davidson’s ability to interact with objects would be like the
solitary creature’s ability to maneuver through the world via non-rational cognitive
processes. Bolted-down Davidson’s inability to determine the distance to objects and
give content to the question of where they are is like the lone being’s inability to give
content to the notion that the objects with which it interacts exist in an objective reality
that stands in contrast to the being’s own subjective beliefs.

The last point is most directly relevant to Davidson’s goal of finishing his
argument against non-linguistic thinkers, for it promises to render grasp of the subjective-
objective contrast dependent on the contrast between each creature’s responses to the
object when viewed against the backdrop of a shared standard. One must be able to make
sense of objective truth in order to make sense of subjective belief. To view reality in
terms of a subjective-objective contrast involves a boundary imposition, a demarcation of
reality into what is internal to self and what is not. The objective world and the
subjective self must figuratively be distanced from each other; the items of the objective
world must be placed in a space outside the self, similar to the manner in which the third
point is literally distanced from the accessible ones in the geometrical case.
Davidson’s idea seems to be that sharing a standard of truth with a second person---a standard that either can fail to meet---opens up the shared space in which objects can be placed outside oneself, and simultaneously gives sense to the notion of subjective beliefs that take these public items as their objects. He wants to claim that recognition of the self as a subject only occurs as one comes to treat another as a subject, and that the need to do so only arises when the interaction between the two beings generates recognition of a shared standard of truth.

Here there are echoes of “Thought and Talk”, and the claim that the shared standard of truth is introduced in order to make sense of divergences between interpreter and interprète as involving error by at least one of them. If we apply this to the triangulation metaphor, the divergences would be comparable to the distance between the two subjects when measured by the shared standard of truth. The very recognition of one’s own perspective on an objective world is accomplished only by differentiating it from another’s perspective. Only with the two perspectives placed at their differing points are we able to pick out the shared object of thought as an item in a public and extra-subjective world.

Considerations from Davidson’s later essays

This take on the use of the triangulation metaphor at the end of “Rational Animals” derives additional plausibility from Davidson’s descriptions in later essays of
the role played by certain aspects of intersubjective triangulation.\textsuperscript{107} He focuses on aspects of intersubjective triangulation that do not necessarily involve language. He claims that a subject must have a history of engaging in activities that involve these aspects of triangulation if it is to meet two necessary conditions for thought. The conditions are not alone sufficient for thought; they must be supplemented by resources only present when the subject’s interactions also include the linguistic aspects of triangulation.

One of these conditions is what I have called the objectivity requirement, the fact that thoughts are subject to the standard of objective truth. The other has to do with identifying an item in an objective world as the object of a belief, what I have called the extensionality requirement. I will appeal to Davidson’s discussions of these conditions for two purposes. In this section, I will explain how these considerations elucidate the nature of Davidson’s appeal to triangulation at the end of “Rational Animals”. But I also believe that these considerations can be used to pursue another line of argument against non-linguistic thinkers, one that is distinct from any argument in “Thought and Talk” or “Rational Animals”. The new line of argument leads us away from questions about what is required in order to grasp the notion of belief, and toward issues concerning the institution of the norms according to which belief is judged. I will pursue that line in the section following this one.

According to Davidson, one important role played by the presence of more than one creature is to pick out an item in the public world as the cause of a perceptual belief.

\textsuperscript{107} Especially Davidson (1992, 1999b, 2001a, 2001b, 2001c).
Picking out the cause is important, Davidson tells us, because in simple cases of perceptual belief the content of a subject’s belief is to be identified with the cause of that belief.¹⁰⁸ This way of talking is a little odd, since belief contents are propositional in structure, and most objects in the public world are not. It is clear enough from Davidson’s discussions, however, that he intends to use triangulation to identify an item in an objective reality as the object of the being’s belief. According to Davidson, simple cases of perceptual belief are to be given special weight in the interpretive process.¹⁰⁹ This is because “…the contents of our earliest learned and most basic sentences … must be determined by what it is in the world that causes us to hold them true. …and given the close connections between thought and language, analogous remarks go for the contents of the attitudes.”¹¹⁰

The point here is that simple cases in which thought and language are applied to the immediate environment—-as they are in native language acquisition—-anchor the relationship between what is thought and said and the world the thoughts and utterances are about. And according to Davidson, it is the cause of our holding a simple belief true that determines the belief’s object. Without triangulation, though, the notion of a belief’s cause is not specific enough to allow a unique object to be identified. Objects are not simply encountered in a void; the world comes at one all at once, and not one unit at a time. I have experiences, and what I have experienced causes beliefs to form. But how is

one to identify, within the sprawl of experience that has led to belief generation, a particular object that is the content-relevant part of the total cause?

By looking at multiple cases in which a being responds similarly, we can hope to identify the aspects of the total situation that are salient to the being. But even then, Davidson argues, our work will not be done. Because thoughts are supposed to have contents about our world, we would like to identify the cause of our subjects’ beliefs with a distal object in a public world. But the behavioral instances we consider will resemble one another not just in one distal aspect, but also in more proximal aspects, as well. We can easily select a distal object that is salient to ourselves across situations, but what basis can we have for identifying that distal object as the one that is salient to our subject across these situations?

Davidson claims that triangulation alone can provide such a basis: “If we consider a single creature by itself, its responses, no matter how complex, cannot show that it is reacting to, or thinking about, events a certain distance away rather than, say, on its skin. The solipsist’s world can be any size.”111 When two beings are responding to the same situations, however, the lines of the triangle can be drawn from the responses of each out into a public realm. Where the lines converge, the triangle is completed at its third apex—a distal object in a shared, public space. This gives an interpreter a principled way to pick out a distal item in an objective world as the content-relevant cause. With a public object of thought identified, we can then hope to determine a specific thought content, which will be objectively true or false of our shared world.

So, Davidson has argued that intersubjective triangulation provides something otherwise lacking to the interpreter, a principled way to capture two necessary aspects of thought: extensionality and objectivity. But this does not suffice to warrant thought attribution, Davidson holds, because it does not show that the beings themselves grasp the objective-subjective contrast. Triangular interaction allows for this concept to develop, because each creature can come to respond not only to objects, but also to other creatures’ responses to objects. A creature can, in effect, come to treat others’ behaviors as signs of the presence of a type of object, by observing correlations between the presence of a type of object and a type of behavior. The “in effect” is important; the beings under discussion are non-linguistic, so someone with Davidson’s agenda must hold that the process of noting correlations between objects and behaviors is non-rational. (This, incidentally, answers Davidson’s aforementioned worry about making sense of a non-linguistic being as having conditional beliefs. If the creature is going to, in effect, treat one phenomenon as a sign of another, this provides a ground for claiming that the creature treats the two phenomena as standing in some form of conditional relationship.) But we can suppose that the task might be accomplished purely through association, which need not be construed as a rational process.

Given a pattern of responses not just to objects, but to others’ responses to objects, “[two beings] A and B are now in a position to notice occasions on which the responses of the other differ from the other’s usual responses … A’s responses are not taken by B to be the same as A’s earlier responses, while B’s responses are the same as
So, intersubjective triangulation of this kind is supposed to be a necessary precursor to a being’s development of the subjective-objective contrast because it alone provides the framework in which a being could come to need to posit such a distinction. Here, a theme from “Thought and Talk” that was under-emphasized in “Rational Animals” reemerges: the distinction between subjective and objective is only made intelligible in contexts that involve error. So, combining the insights of several of Davidson’s discussions, we have something like this:

1 Only intersubjective triangulation presents a being with divergences from an established practice.
2 A being must be presented with divergences from an established practice if it is to make sense of the notion of error.
3 The notion of belief is introduced and initially applied in order to account for a specific kind of error: divergence from the standard of truth.
4 So, only beings who have engaged in intersubjective triangulation can be in a position that renders the notion of belief understandable.

So far, the argument falls short of establishing that only language users can have the concept of belief. So it cannot complete the master argument of “Rational Animals” as it stands. Davidson has argued that only intersubjective triangulators can have the concept of belief, but non-linguistic beings are capable of intersubjective triangulation. Davidson needs an argument that only those who engage in a specifically linguistic type of triangulation can understand the notion of belief.

Here, Davidson’s idea seems to be that the other being must be recognized as a subject in order for the idea of a shared standard to make sense. If a being simply

---

correlates responses of another being with an event, this does not show any recognition that the other being has a point of view. The other being could simply be treated as an additional object, perhaps one that is useful as a sign of various other things, but merely an object nonetheless. All of this is compatible with triangulation, but it does not yet yield a notion of error, because it does not yet force the view that the other being is a subject. For all that has been said so far, the other being’s occasional failures to move in the anticipated way could be treated merely as reasons to believe its movements are not perfect signs of the correlated features, and not as reasons to believe it is representing an objective world incorrectly. So, it does not provide either being with any reason to come to think of itself and the other being as subjects, and as such capable of being subject to a shared standard.

According to Davidson, “Only communication can provide the concept [of objectivity], for to have the concept of objectivity, the concepts of objects and events that occupy a shared world, of objects and events whose properties and existence is independent of our thought, requires that we are aware of the fact that we share thoughts and a world with others.”¹¹³ So, my idea that I am a subject whose thoughts represent an objective world of public items dawns along with the notion that someone else is a subject whose thoughts represent an objective world of public items. And “for two people to know of each other that they are so related, that their thoughts are so related, requires that they be in communication.”¹¹⁴ So, Davidson’s idea is that only by communicating, presumably linguistically, with something else can a being come to see

¹¹⁴ Davidson (1992), p. 121.
that this something else is also someone else, and only by seeing that there is such a thing as someone else do we come to understand the notion of a thinking subject at all.

So, it is not enough that a being triangulates; it must also be aware that triangulation is intersubjective, and only language, which cannot help but have a shared standard of truth, forces the idea of intersubjectivity upon us. We can rewrite our new Davidsonian line to reflect these ideas:

Davidsonian Master Argument Subargument 2, version 2:

1 Only a being that recognizes triangulation to be an intersubjective practice governed by a shared standard will be able to understand divergences from expectations as instances of error.
2 A being must understand divergences from expectations as instances of error if it is to make sense of the notion of belief.
3 Only linguistic communication causes a being to understand that it shares a standard—intersubjective truth—with others.
4 So, only linguistic communication suffices to bring a being to understanding of the notion of belief.

This line is capable of completing the master argument of “Rational Animals”. Annexed to any of the three versions of Subargument 1, we get a detailed defense of the claim that only language users can be thinkers, via the subconclusions “Only beings that grasp the concept of belief can be thinkers” and “Only linguistic beings can grasp the concept of belief.”

The Institution of Norms in the Social Practice of Language
One might employ some of the insights from Davidson’s discussion differently. Consider again the basic line on which Subargument 2, version 2 was based:

1 Only intersubjective triangulation presents a being with divergences from an established practice.
2 A being must be presented with divergences from an established practice if it is to make sense of the notion of error.  
3 The notion of belief is introduced and initially applied in order to account for a specific kind of error: divergence from the standard of truth.
4 So, only beings who have engaged in intersubjective triangulation can be in a position that renders the notion of belief understandable.

Davidson promotes the idea that a subject’s participation in a social practice makes an indispensable contribution to the subject’s status as a thinker. But he does so by focusing on the role the practice plays in giving the subject reflective capacities. This strategy makes sense for Davidson, since he also argues that thinkers must have reflective capacities.

But suppose one is not convinced that thought requires reflection. One might still attempt to exploit the fact that language is a social practice in order to argue against languageless thinkers. For even if thinkers need not have the capacity to reflect on their own minds, they clearly do have to be properly held accountable to the standards of rationality and truth. And it might be argued that standards such as these require institution in a practice with certain features, and that these features are found only in the practice of language use. Perhaps it is only appropriate to apply such standards to those who participate in linguistic practice.
Consider a situation in which we encounter an unfamiliar being. Obviously, the newly encountered being should be treated as a physical system, bumped about in accordance with the laws of physics, as all physical systems are. But is the new being a thinker? We can suppose that the being moves in ways that are amenable to explanation in terms of reasons. When we give such explanations of its movements, we are judging it by rational standards. To justify the claim that the subject is a thinker, we need to show not only that we can gain an interesting account of the being’s movements by applying rational standards. We also need to show that those standards can properly be considered to govern the being’s movements. If not, we are applying our standards outside of their true domain, and though this might yield valuable results, it will not change the fact that the being is not really subject to the standards we have imposed. The being’s actions will be describable in terms of rules, but it will not really be subject to those rules.

What can show that it is appropriate to apply a standard to a being? The being’s explicit avowal that it is guided by the standard in its actions will do, as will its signature on a contract that specifies an obligation to meet the standard. Clearly, non-linguistic beings are unable to supply either of these. But often we go without them in the case of linguistic beings, as well. Thus, we have the notion of implicit consent, in which beings in effect take on various obligations despite never explicitly accepting them. Typically, initiation into a norm-governed practice suffices to bind a being by that practice’s norms, even if the being is not aware of all the norms and cannot claim to have mastered the rules of the practice. Tyler Burge’s character who believes that arthritis can cause pain mid-femur has not mastered the rules for application of the concept ARTHRITIS, but
because he has been initiated into a linguistic practice in which that concept is mobilized by the term “arthritis”, his utterances of “arthritis” are nonetheless subject to the normal rules for the term’s use.115

So, initiation into a practice carries much weight when it comes to our judgments of whether a being is appropriately held accountable to norms. From these considerations, one can generate a simple argument against non-linguistic thinkers:

**Basic Institutional Argument:**

1. Thinkers’ behavioral and cognitive processes must be accountable to standards of truth, rationality, and concept application, not merely describable in terms of such standards.
2. Only beings who have been initiated into a practice are accountable to that practice’s standards (in this case, standards of truth, rationality, and concept application).
3. Only linguistic beings have been initiated into a practice governed by standards of truth, rationality, and concept application.
4. So, only linguistic beings are thinkers.

**The Kripkenstein Challenge**

The nature of the practices in which a subject participates also impacts which norms we are going to hold the subject accountable to. The counterpart to Burge’s character, who lives in a society where “arthritis” is applied to ailments of the bones as well as joints, is not accountable to the rules of application for ARTHRITIS, but to those of the distinct concept to which his society applies “arthritis”. This is the case despite the fact that the counterpart is supposed to be internally identical to the original character. So, one might suspect that a being’s participation in a social practice in which linguistic

---

terms are manipulated according to rules is necessary in order for there to be any fact of the matter as to which norms the being is accountable to.

This is reminiscent of Davidson’s worries about our inability to determine which concepts a non-linguistic being might employ. But Davidson portrays this as worrisome mainly because we are accustomed to making fine-grained attributions to linguistic beings in order to whittle down the range of interpretive options. But now, our worry is not just that uncomfortably many interpretive options will be left open. It is also that our inability to discriminate between different patterns of attributions that fit the subject’s behaviors indicates a failure for there to be any fact of the matter as to which rules we ought to hold the being accountable to. The worry is that of Kripke’s Wittgenstein (or, as I shall refer to him, Kripkenstein)\textsuperscript{116}: no matter what a subject does next, there is a rule according to which its action fits with what it has done in the past. So, without a basis for choosing a rule as the one that our subject is accountable to, “there would be neither accord nor conflict here.”\textsuperscript{117} Insofar as we fail to find either accord or conflict in a subject’s behavior, we fail to make sense of the subject’s behavior as normatively constrained. If we cannot find a principled way to hold a subject accountable to a rule for a particular concept’s application, our very ability to understand the subject’s movements in normative terms will be threatened.

Kripkenstein develops his worry by use of a mathematical example involving the “+” sign.\textsuperscript{118} The rule for use of “+” ought to determine a proper answer for indefinitely

\textsuperscript{116}See Kripke (1984), p. 7.
\textsuperscript{117}See Wittgenstein (1958), section 201.
\textsuperscript{118}Kripke (1984), pp. 7-9.
many as yet unencountered math problems, but what fact determines which rule one actually follows in one’s use of “+”? We can imagine a skeptic who exploits the fact that, for each of us, there is some number n such that we have never performed a calculation involving “+” with any number as great as n. The skeptic defines “quus” as a function that operates exactly like the familiar plus function when all inputs to it are less than n, but gives the answer 5 when any input to it is n or greater. The skeptic then challenges us to find any fact that might determine which of the rules we have followed in the past, and hence, whether we ought to answer a new problem involving “+” and an input greater than n with a sum or the “quum” of 5 in order to keep following the same rule we have in the past.

As Kripke indicates, the reference to the rule followed in the past is not an essential part of the puzzle. The skeptic’s real target is the claim that there is a fact of the matter as to what rule one is following now, in any given act of supposed rule following, regardless of what one has done in the past. The skeptical challenge to the claim that there is a fact of the matter as to meaning ‘plus’ or ‘quus’ by one’s use of “+” extends to a challenge that there is a fact of the matter about what anyone means by anything by the use of any term.

Kripke argues at length against solutions to the challenge that are not based on appeal to community practice. Several of these---involving the subject’s explicit self-instruction to add, computation according to the same rule as before, or having learned an algorithm for addition---appeal to question-begging resources. Other replies, involving

---

120 Ibid. pp. 15-54.
introspectively accessed phenomenal states of the subject, are rejected because those states cannot attain the epistemic status needed to justify either the claim that the subject was adding or the claim that the subject was quadding. Strategies that appeal to the subject’s dispositions face the difficulty that we are disposed to make mistakes, so the rule the subject is accountable to will not be identical to the rule that perfectly fits the actions the subject is disposed to perform. One can try to solve this problem by considering an idealized version of the subject, but what are we supposed to be idealizing away from: situations in which the subject’s behaviors do not fit with addition, or situations in which the behaviors do not fit with quaddition? The skeptical challenger will view either answer as question-begging. Finally, appeals to simplicity are supposed to fail because simplicity considerations are relevant only as an epistemic crutch when we lack access to some facts that would establish the truth.

I am not convinced by Kripke’s argument against simplicity considerations, and I will appeal to them in my own account in Chapter 4. The dispositional response deserves further consideration, as well---perhaps the dispositions of an idealized subject to count certain moves as errors and others as correct when considering its own past behaviors could help to answer the challenge. But Kripke takes himself to have eliminated all plausible solutions that fail to look beyond the subject to community practices.

With a community practice in place, however, we can readily make sense of the subject as accountable to one rule and not the other. The fact that multiple subjects come together to engage in one shared practice implies a sharing of standards. By engaging in the group practice, a being takes on the practice’s standards. Thus, Burge’s character is
accountable to his community’s standards when he uses “arthritis”, and Kripke’s character is accountable to his community’s standards when he uses “+”.

The Kripkensteinian proposal is to take the interpretive practice of a community at face value. Community practices of acceptance and censure embody the conditions under which it is permissible to attribute various concepts to subjects. Since the community carries on with the practice, it must play some role in their lives that they find worthwhile, and this is all the justification it needs in order to be accepted.

One objection to this solution is that the skeptic can raise the same kind of challenge with respect to community rules. The skeptic renders plausible the claim that there are infinitely many gerrymandered concepts that fit a subject’s behavioral dispositions by pointing to examples of how one might generate such concepts. But this method is no less available with respect to communities than it is with respect to individuals. The “+” example involves choosing a number greater than those that one subject has ever taken as an input to the function labeled “+”, and showing how “+” could be interpreted equally well as meaning plus or quus. But then choosing a number greater than those that any member of the community has ever taken as an input to the “+” function ought to have the same impact on the matter of determining which rule the community is really reinforcing, and which ones it is censuring. If patterns of past usage, dispositions to use a term, etc., are insufficient to insulate a lone subject’s behavior from an indeterminacy of norms, then they will also fail in the case of a community.

Though Kripke does not discuss this objection, it can be answered by noting the ways in which interpretive authority is dispersed among community members. There
will always be new instances to which no one in the community has had an opportunity to attempt to apply concepts. But linguistic practice treats concepts as objectively applicable to as yet unencountered cases. Most of the time, there is little question of how to continue; practitioners deal with new cases automatically and without explicitly considering different options. But when disputes arise, linguistic practice offers ways of settling them, and that is key to answering the skeptical challenge by appeal to community. If two practitioners disagree about whether to apply a familiar term to an unfamiliar instance, they can offer reasons why their usage should be preferred. Furthermore, as Putnam pointed out, there are some within communities who hold the status of expert with respect to various terms and their applications. In cases where there is a consensus among experts, such individuals are treated as understanding how to properly apply the terms to new instances as well as familiar ones.

The most important point, though, is not about the ways in which disputes can be settled within interpretive practice. It is rather the fact that there are ways of settling disputes, and that these are treated as capturing facts about how the concepts should have been applied all along, and not as de jure pronouncements over matters that were hitherto indeterminate. We have genuine accord or conflict in cases where a subject has applied a concept to a new instance because the subject is a participant in a practice in which there are ways of determining accord and conflict that are treated as revealing objective truths. The dispersal of authority shows that not just any move an individual makes can be counted as correct, and the fact that the standards that determine correctness and

---

121 See Putnam (1975).
incorrectness are treated as objective shows that the skeptic cannot merely repeat the argument at the level of the community by replacing the whims of the individual to count an action as correct with the whims of the community.

So, there is a way to solve the Kripkenstein problem by appeal to linguistic practice. Kripke has not devastated the very notion that beings are accountable to standards for concept application. But he has argued that there is no solution apart from considerations of linguistic practice. So, one can generate the following argument against non-linguistic thinkers:

*Kripkensteinian Argument:*

1. Thinkers’ behavioral and cognitive processes must be constrained by standards of concept application.
2. Considered in isolation from linguistic practices, the activities of a subject fail to determine whether the subject is accountable to a standard of concept application that it meets, or to a standard it fails to meet.
3. If it is indeterminate whether a subject’s activities are in accord with the standard of concept application to which it is accountable or in conflict with the standard to which it is accountable, then the subject is not constrained by standards of concept application.
4. If a being was not initiated into linguistic practice, then it must be considered in isolation from linguistic practices.
5. If a being is non-linguistic, then it was not initiated into linguistic practice.
6. So, if a being is non-linguistic, then it is not a thinker.

Interestingly, Kripke himself does not advocate this argument. He seems to reject, or at least call into question, the premise at line 4. Kripke holds that, “Our community can assert of any individual that he follows a rule if he passes the test for rule
following applied to any member of the community.”⁵¹²² So, as long as a linguistic community has tests for concept application that can be passed without actually using language, the community can assert that the subject who passes the tests follows the rule.

Whether this counts as rejection of line 4 or not is unclear. Obviously, it depends on whether Kripke thinks that linguistic communities have tests for concept application that can be passed without language use. Like Tyler Burge,¹²³ I am inclined to think they do, despite Davidson’s worries about how exactly such tests would be performed. But it also depends on whether or not Kripke takes the permissibility of the community’s assertion that the subject follows the rule to indicate that the subject really is accountable to the rule. One might hold that there is nothing more to a subject’s being accountable to a rule than that a community can permissibly hold it accountable to the rule. That would deny 4, because it would license consideration of non-linguistic beings according to standards of linguistic practice.

On the other hand, one might hold that the community’s assertion could be permissible even though the subject might turn out to not really follow the rule. That still seems to deny 4, if one takes 4 out of context, because it allows the community to consider the subject as following the rule. In context, however, 4 should be interpreted as meaning “If a being was not initiated into linguistic practice, it must be considered in isolation from linguistic practices for the purpose of determining whether it is really accountable to standards of concept application.” The proposal we are considering licenses the community to consider the subject according to standards of linguistic practice.

¹²² See p. 110.
¹²³ See Burge (1979).
practice, but only for the purpose of generating warrant to describe the subject’s actions in terms of rule-following, not for the purpose of determining whether the creature really is accountable to the rules. So, this proposal accepts 4 after all.

I would expect a Kripkensteinian to complain that no good sense can be made of the distinction I have just drawn between purposes. I would expect a Kripkensteinian to advocate my earlier proposal that obtaining warrant to describe a subject’s actions in terms of rule-following is determination that the subject really follows rules. If I am right about this, a Kripkensteinian would not accept line 4 on the basis of a distinction between the purpose of obtaining warrant for asserting that a subject follows a rule and the purpose of determining that the subject really does follow the rule.

But a Kripkensteinian could still accept line 4, and hence the Kripkensteinian Argument, on other grounds. The Kripkensteinian could reject the idea that there are non-linguistic tests that suffice to license assertions that a being applies concepts. Alternatively, one could still stand firmly within the Kripkensteinian camp and yet reject Kripke’s claim that a community can assert of any individual that he follows a rule if he passes the test for rule following applied to any member of the community. One might hold instead that such tests are designed for community members, all of whom have been initiated into the practice. Because of the shared background of those for whom the tests were designed, the tests need not control for as many potential hidden variables—all of these will be constants within the community due to their shared background. But an outsider brings a different background to the test situation, one the test has not been designed to handle. So, when it comes to testing for concept application, more might
reasonably be required of a non-member, about whom we cannot assume as much in terms of background conditions.

**Brandom’s Institutional Considerations**

Two of the key considerations I invoked in defense of a community-based solution to the Kripkenstein puzzle were the dispersal of authority among community members and the fact that the standards for correct action are treated as objective by the practitioners. These two factors figure prominently in Robert Brandom’s arguments that only language users can be thinkers. Despite this fact, Brandom’s approach to issues of thought and language has more in common with Davidson’s approach than Kripkenstein’s. All three appeal to linguistic practices in order to explain thought. But whereas Kripkenstein wants to take attributive practices at face value, Davidson argues that thought and language must arise simultaneously, because several key aspects of both thought and language make sense only with the interplay between the two in place.

Brandom is closer to Davidson’s camp, because he also emphasizes that several key aspects of thought arise within a community because of their relationships to aspects of language use in that community. Thoughts must have contents, and for Brandom semantic notions such as content are theoretical abstractions from the proprieties of linguistic practice. Like Davidson, Brandom demands a degree of meta-awareness on the part of thinkers. It is not enough that a being responds to the world in ways that in fact fit a practice’s standards; the being must be to some extent aware that it is participating in a
practice. And it is only when practitioners are aware enough of the practice itself to make explicit the rules they have been following that they will even come to grasp semantic notions. The ability to attribute thoughts to oneself is derived from the social skill of attributing thoughts to others.

Furthermore, thought contents must be objectively true or false, and for Brandom this type of normativity can be generated only by the relationships of authority involved in a social practice in which linguistic tokens are manipulated. For Brandom, as for Davidson, the key point is not that the practice generates standards that can then be applied to anyone (as in the Kripkensteinian account), but that linguistic interactions occur between persons. It is the interplay between the first-person and the second-person perspectives that allows for the mandatory dispersal of authority, and forces each interlocutor to acknowledge a standard that goes beyond the whims of either.

Because of these similarities, Brandom’s arguments are often structured in ways that are reminiscent of Davidson’s arguments. In fact, Brandom claims at one point to be filling in the details of the master argument schema from “Rational Animals”.\textsuperscript{124} Hence, in Chapter 4, I will appeal to some of Brandom’s discussions in my assessment of the premises of Davidsonian arguments against non-linguistic thinkers. But there are two distinct arguments against non-linguistic thinkers that can be pulled from Brandom’s work.

One of these lines of argument focuses on the fact that authority within a practice typically must be earned. Because thought attributions hold subjects accountable to

\textsuperscript{124} See Brandom (1994), p. 152.
rational standards, attributers treat subjects as having authority over their behaviors and cognitive processes. Rational standards are derived from the standards implicit in linguistic practice, on Brandom’s view. Perhaps then authority over one’s own behaviors and cognitive processes can only be earned by developing some degree of mastery over linguistic practice. Brandom holds that deliberation requires proficiency at the social skill of scorekeeping in the game of reasons: “…in deliberating, an agent considers what commitments would be attributed by scorekeepers, under various circumstances. It is for this reason one must be able to assess the conduct of others in order to deliberate about one’s own.”\footnote{Brandom (1994), p. 270.} If deliberation is required for authority over one’s behaviors and cognitive processes, then one can complete an argument against non-linguistic thinkers.

\textit{Brandomian Deliberation Argument:}

1 Deliberation involves considering what commitments would be attributed by linguistic scorekeepers under various conditions.
2 Only beings who have become proficient at the skill of linguistic scorekeeping can consider what commitments would be attributed by linguistic scorekeepers under various conditions.
3 Only linguistic beings are proficient at the skill of linguistic scorekeeping.
4 Deliberation is required for authority over one’s own behaviors and cognitive processes.
5 To be accountable to rational standards, beings must have cognitive authority over their behaviors and cognitive processes.
6 Thinkers as such are accountable to rational standards.
7 So, only linguistic beings can be thinkers.

The other Brandomian argument focuses on interpersonal relationships of authority and interpersonal substitutions of terms. For Brandom, semantic content is a theoretical abstraction from the pattern of moves that are allowed, disallowed, and
demanded within the “game of reasons”, the social practice in which justifications are requested and offered. Performances---utterances---are the primary bearers of content, and an utterance’s content is determined by the deontic statuses it confers upon the utterer and the audience. The content of the utterance is thus determined by its relationships to other utterances that are made permissible or impermissible by the overall pattern of utterances that have been made in the course of conversation.

So, for Brandom, content is determined by inferential relationships, not by relationships between a representational state and items in the world. Nonetheless, Brandom concedes that content has a representational dimension,\(^\text{126}\) and he explains this in terms of interpersonal inferential substitutions involving subject and predicate terms. Inferences involving subject and predicate terms are important because Brandom takes the object-property structure of the world to be determined by the subject-predicate structure of language.\(^\text{127}\) It is the presence of subject and predicate terms that allows a being to confront a world of propertied objects, as it must if it is to have states that bear conceptual contents. Subject-predicate structure is itself abstracted out of sentential structure, which is abstracted from norms for sentence use, which are implicit in the linguistic practice of scorekeeping in the game of reasons. So, linguistic practice is needed in order to make sense of a subject as thinking about propertied objects.

So, we have the following argument:

*Brandomian Syntactic Priority Argument:*

\(^{126}\) Ibid., p. 431.  
\(^{127}\) Ibid., see Chapter 6.
1 Thoughts must bear conceptual contents that represent a world of objects as having certain properties (i.e., to be a thinker, one must confront a world of propertied objects).
2 Only a being who participates in a linguistic practice with subject-predicate structure confronts a world of propertied objects.
3 So, only a being who participates in a linguistic practice with subject-predicate structure is capable of thought.

In fairness to Brandom, it should be noted that he sometimes writes as though it is only a special form of thought that he denies to non-linguistic beings. Thus, we are told that non-linguistic beings can still have a derived form of intentionality, and that they can still have beliefs in a parasitic fashion. I believe these concessions should not be made. Brandom’s arguments, if successful, show that attitudes cannot seriously be attributed to non-linguistic beings. On Brandom’s view, non-linguistic beings are not practitioners of the game of reasons, so on Brandom’s view they are not genuinely accountable to the norms that bind thinkers. Brandom might be reluctant to draw a firm line between thinkers and non-thinkers because on his view, a subject’s status as accountable to norms is not supposed to be reducible to any set of facts about the subject. Thus, it might seem inappropriate for Brandom to give a theory about who is and who is not accountable to norms; that matter will only be decided by those who actually choose to hold the being accountable to norms or not. But Brandom has clearly established some necessary conditions a being must meet in order for it to be properly held accountable to rational norms. And, if his arguments are successful, non-linguistic beings cannot meet these conditions.

128 Ibid., p. 143.
Conclusion

At the outset of this chapter, I identified seven purported requirements on interpreting a subject as a thinker. These were the requirements of:

- control
- reflection
- conceptual standards
- intensionality
- extensionality
- objectivity
- institution

Over the course of the chapter, I have drawn from the work of interpretationists ten arguments against the possibility of non-linguistic thinkers. Each of these arguments exploits at least one of the seven purported requirements. In Chapter 4, I will defend the possibility of non-linguistic thinkers against these arguments by showing that non-linguistic beings can meet all the genuine requirements for thought. For some purported requirements---control, concept application, intensionality, extensionality, objectivity, and perhaps institution---I will show how a non-linguistic being can meet them. For others---reflection and perhaps institution---I will argue that a weaker requirement suffices for thought attribution, and that non-linguistic beings can meet the weaker requirement. I will show how my set of requirements respects the intuitions that drive the ten arguments; yet I will also show how each of the ten arguments fails to establish the

129 See Appendix to this dissertation for a list of the ten arguments, complete with indications of the points at which each exploits one of the seven purported requirements for thought, and explanations of the derivations' structures.
impossibility of non-linguistic thinkers, once the requirements on thought are understood in the manner I defend.
Chapter Four: The Failure of Arguments from the Requirements to Linguistic Priority

Introduction

In Chapter 3, I identified ten arguments against the possibility of non-linguistic thinkers. Each of these arguments relied in some way on one of seven proposed requirements for a subject to count as a thinker. In this chapter, I will argue that a being could be non-linguistic and yet meet every genuine requirement for counting as a thinker.

This means that I must argue against at least one premise from each of the ten arguments from Chapter 3. But my method will not be merely destructive. I will show how one can respect the intuitions that motivated those arguments without denying the possibility of non-linguistic thinkers. Some of the seven proposed requirements---control, standards for concept application, intensionality, extensionality, objectivity, and perhaps institution---are genuine. Others---reflection and perhaps institution---are misconceived. For each genuine requirement, I will show how a non-linguistic being could meet it. As a result, some of the ten arguments will be shown to misapply the requirement in a premise. Other arguments correctly apply the requirement, but contain another premise that is too restrictive given the ways in which the requirement can be met. For each misconceived requirement, I will show how the motivations that lead to
advocacy of that requirement actually support a different requirement, one that a non-linguistic being could meet. I will show how the arguments against non-linguistic thought that relied on the misconceived requirements fail when altered to respect the actual requirements. The result of this process will be the outline of a view of the thinking mind that draws a principled distinction between genuine thinkers and beings whose actions are merely describable in interpretive language, but does not deny the possibility of non-linguistic thinkers. I will elaborate upon the view and its application in Chapter 5.

The Intensionality Requirement

I will begin with the intensionality requirement. If we are to hold a being accountable to rational standards, we must be able to identify not only the objects of its various attitudes, but also the manner in which it thinks about them. A subject who wishes to see the Morning Star while avoiding seeing the Evening Star is not necessarily irrational, despite the fact that the subject wishes for the impossible---to see Venus while avoiding seeing Venus. By contrast, a subject who wishes to see the Morning Star while avoiding seeing the Morning Star is plainly irrational, though again this is a wish to see Venus while avoiding seeing Venus. The key difference between these two subjects cannot be found in the objects of their thoughts---both are thinking about Venus. The important difference is in the way in which each thinks about Venus. The differences in
the specific concepts each employs in order to think about Venus yield two different verdicts about their desires. One is clearly irrational; the other is not.

Thus, judgments of rationality or irrationality are sensitive to the ways in which subjects think of objects. This motivates a requirement that an interpreter be able to identify the manner in which a being thinks of objects if the interpreter is to treat the being as a thinking subject. This requirement plays an interesting role as a premise in three of the ten arguments for linguistic priority: the Davidsonian Network, Holism, and Generality arguments.130

**Intensionality and the Network Argument**

I will begin by considering how the intensionality requirement is employed in the Network Argument. The Network Argument has the following structure.

1 The specific conceptual composition of a thought is determined by its place in a dense, extensive network of beliefs.
2 Only linguistic subjects behave in ways rich enough to embody a dense, extensive network of beliefs.
3 So, only linguistic subjects can have thoughts with determinate conceptual contents.

The Network Argument relies on the intensionality requirement in order to motivate the premise at line 1. The demand for specific conceptual composition is a demand that we capture the manner in which a being thinks about an object in order to

---

130 It also appears as a premise in the Brandomian Syntactic Priority Argument, but it is not the main source of interest and difficulty in that argument. The difficult issues with the Syntactic Priority Argument are raised by the Extensionality requirement, so I will discuss the Syntactic Priority Argument in that context.
treat the being as thinking about the object at all. Line 1 contains the additional claim that a thought’s specific conceptual content is determined only by the relations between that thought and others in a dense, extensive network.

The premise at line 2 is motivated by Davidson’s considerations involving the dog who behaves in an excited manner upon its master’s arrival: “The dog, we say, knows that his master is home. But does it know that Mr. Smith (who is his master), or the president of the bank (who is that same master), is home? We have no real idea how to settle, or make sense of, these questions.”131 A being that has attained linguistic status is guaranteed to have at its disposal a network of concepts as dense and extensive as those of a public language. Because of this, we can work toward an understanding of how a linguistic being thinks of various objects by paying attention to the words the subject chooses to employ in order to talk about those objects. But for a non-linguistic subject, Davidson claims that we will have no way to make sense of the subject as employing one concept rather than another in a given cognitive episode.

I have no quarrel with the manner in which the intensionality requirement is employed in the Network Argument. The demand for specific conceptual composition is legitimate. And I have no quarrel with the claim that a thought’s content is determined, at least in part, by its relations to other thoughts, though it will be clear from my discussions in chapters four and five that world-to-mind relations generated by perception and mind-to-world relations embodied in actions should be included in the network as well. Furthermore, I will grant that linguistic capacities increase the density of the

---

131 Davidson (1975), p.163.
connections in a subject’s network of beliefs. As a matter of epistemology, this makes it much easier for an interpreter to discern the relations the thought bears to others and thus allows for greater precision in the practical task of offering a working interpretation of a subject’s thoughts.

But I dispute Davidson’s claim that a lack of language use on the part of a subject prevents us from being able to make sense of questions that ask how the subject thinks of an object. And I dispute the associated claim made at line 2 of the Network Argument: that non-linguistic beings do not behave in rich enough ways for their cognitive systems to instantiate a network of thoughts sufficiently dense to allow for holistic content determination. I will argue, on the contrary, that a lower web density suffices for content determination.° In the interest of clarity, it is worth stating that this sufficiency claim is made only about the intensional aspect of thought. A lower web density will suffice to answer concerns about the possibility of non-linguistic thinkers insofar as, and only insofar as, those concerns are based on the requirement that thoughts be composed of specific conceptual contents.

As a preliminary to my point regarding web density, I would like to return to a claim I made when I discussed the dog-and-master scenario in Chapter 3.°° There, I stated that we should not say that the dog thinks the president of the bank is home, despite the fact that the dog’s master is the president of the bank. Davidson claims that we cannot even make sense of the question, “Does the dog know that the president of the bank is home?” On the contrary, we can make good enough sense of that question to

---

132 I thank Eric Carter for suggesting this approach to me.
133 See the section “Considerations Davidson Marshals in ‘Thought and Talk’”.
143
know that the answer is “no”. (The fact that Davidson discusses a particular kind of belief, knowledge, does not seem to play an important role in the argument, so I will discuss conditions for belief attribution in what follows.)

We can begin by considering the way we would go about answering that question for a linguistic subject who knows that the dog’s master is home. Suppose we are aware that our imagined linguistic subject does not know that the dog’s master is the president of the bank. We would not want to attribute to our subject a belief that the president of the bank is home. We typically rule out any attributions that clearly do not capture the way in which the subject thinks of the object. Suppose further that our imagined linguistic subject has lived a financially sheltered life, and has had neither opportunity to experience objects that are banks nor need to treat objects differently based on whether those objects are banks or not. We certainly would not want to attribute to the subject a belief that the president of the bank is home. That interpretation would attribute a concept, BANK, that our subject has had no opportunity to acquire and that sets a standard for concept application that our subject has no need to meet in its cognitive activities. If the subject uses a word that means “bank”, we might for Burgean reasons be driven to attributions that include BANK, but given that our subject is non-linguistic, the subject’s lack of any interest in and sensitivity to whether or not things are banks is a compelling reason to avoid interpretations that attribute beliefs containing BANK as part of their content.

So, against Davidson, we have good enough reason to insist that the dog does not believe that the president of the bank is home; we can make good enough sense of the
question whether the dog believes that. But another Davidsonian question remains: what content would be appropriately attributed to the dog? Suppose we eliminate all attributions that contain concepts the dog would have neither need nor opportunity to grasp. Should we describe the way it thinks of its master as “master”, “friend”, “food giver”, or any other of the many descriptions that accurately describe Mr. Smith and use only concepts a dog might be taken to possess? Davidson might complain of any such proposed attributions that it is unclear that a dog would possess these concepts, because there are simply not enough other concepts in play for us to identify any concepts at all. If we wish to attribute an understanding of masterhood, that seems to imply that we should also be willing to attribute an understanding of pethood, or at least of some other subservient status. A dog probably would not have reason to understand these roles as we conceive of them, and MASTER simply cannot be identified as such in the absence of its connections to these roles.

Notice that the key issue here is not actually the density of the web surrounding MASTER, though. The key issue is whether certain particularly closely related concepts are in place. MASTER needs PET or SERVANT or something else along these lines. BANK needs CURRENCY or something similar. The president of the bank will presumably have a much denser network embodying his notion of bankhood than will a child who has just opened a first savings account. The child might have very few direct links between beliefs about banks and other beliefs, but as long as the basic connections between BANK and a few other key concepts are in place, the child can think thoughts containing BANK.
We can imagine a representation of an ideal web of belief concerning banks, with various nodes representing various beliefs that contain BANK and other beliefs that do not themselves contain BANK but connect to those that do. We can then imagine eliminating links between nodes until we are left with a representation of the bank president’s web of belief with respect to banks. Though we will have eliminated much of the density of the ideal network, this network will still have a reasonably high density. The fact that the president has thoughts about banks as such will be determined by the pattern made by the links between nodes, along with the relations the nodes bear to various perceptions and actions. Of course, no single link is going to be absolutely necessary for identifying the pattern as determining that the president has thoughts about banks; that is why we could stray from the ideal and still have a web that contains beliefs about banks. But it will be important that certain kinds of conceptual links, such as that between BANK and CURRENCY are supported by the remaining links between beliefs.

Now consider what happens when we pare down further, from the bank president’s web of belief to that of the child. By hypothesis, the child has thoughts about banks. But there will be very little density left in this web. The child has few beliefs about banks and probably a significant number of them involve misconceptions. But as long as a few key conceptual links are preserved by the thoughts about banks, currency, etc. that remain in the depleted and deformed web we are now considering, the low-density web can suffice to determine that the child thinks about banks as such.

Matters might actually be even simpler than this when it comes to establishing that we can attribute thoughts containing BANK to the child. The child is linguistic, and
uses a word for banks. Burgean considerations suggest that we can attribute thoughts involving banks despite the fact that the child has only a poorly-formed notion of what banks are.

But of course, this strategy is not available for a dog. So, if we wish to attribute thoughts to a dog, we will need to restrict our selection of concepts to ones that depend for their identification only on other concepts a dog might plausibly be taken to possess. Concepts like FOOD, WATER, HOME, etc. that identify features of fundamental importance to a dog’s livelihood and that the dog frequently encounters in objects it experiences would be of the right kind to serve in attributions to the dog. As long as there is a link between FOOD and EAT, between WATER and DRINK, between HOME and DEFEND, even a sparse network could suffice to determine that the dog thinks of these things as such.

Can we safely say then that the dog believes that the food giver is home? Insofar as we are worried about whether the key conceptual links could be in place between the attributed concepts and those concepts that give them shape, my only concern would be over GIVER. It perhaps has connections with PROPERTY and AGENT that would make it inappropriate to attribute to a dog. The most appropriate attribution I can think of plays on the fact that dogs are pack animals, and as such must be able to recognize and respond appropriately to statuses in a social hierarchy. The dog believes the pack leader is home.

Again, I want to remind the reader that I do not claim to have secured the right to make this attribution at this point. But I believe I have defended it against the specific
charge made by the Network Argument. Line 1 of the Network Argument contains a false premise. High density is not actually required in order for a web of beliefs to determine a specific conceptual content for a particular belief. A small number of links between beliefs containing that concept and beliefs containing importantly related concepts can suffice. Thus, though high density does depend on language use in the manner Davidson suggests, the intensionality requirement can be met in the absence of high density. So the Network Argument fails to establish that only language users can meet the intensionality requirement.

*Intensionality and the Holism Argument*

Davidson has two other intensionality-based arguments against non-linguistic thinkers, however. One of these is the Holism Argument.

1. If Ramseyan methods cannot be applied, then the holism of the mental leaves open an unacceptably large number of competing interpretations.
2. If we cannot judge subjects’ selections at the grain of intensions, then we cannot apply Ramseyan methods.
3. If a subject is non-linguistic, then we cannot judge the subject’s selections at the grain of intensions.
4. So, if a subject is non-linguistic, then the holism of the mental leaves open an unacceptably large number of competing interpretations.

The Holism Argument provides an independent motivation for the intensionality requirement. A subject’s ability to meet the intensionality requirement is treated as necessary in order for an interpreter to discern the preferences that are implicit in the
subject’s pattern of selections. In order to determine a subject’s preferences, we need to know not only which items the being favors, but also which features of those items lead the being to favor them.

Many patterns of belief-desire attributions will make reasonably good sense of a given pattern of behavior; this fact is captured by line 1. Davidson believes that the existence of a decision-theoretical method for determining a being’s preferences allows us to make sense of one of the many potential belief-desire assignments as the one that correctly describes the subject; paired with the intensionality requirement, this yields line 2. Line 3 claims that only linguistic behavior allows a being to satisfy the intensionality requirement with respect to determination of the being’s preferences.

I will make a concession to Davidson here. I think he is right that a subject’s lack of linguistic behavior threatens our ability to discern in the subject’s behavior a clear, unique pattern of preferences. Considering one of Davidson’s examples, suppose a subject is presented with an apple and a pear and it is clear that the subject can have exactly one of the two. If the subject chooses the apple, does this show a preference for apples over pears, for red things over yellow things, for items that will be good to eat three days from now rather than at the present moment, or what? One might think that the problem could be solved by varying some features while keeping others stable, until a pattern emerges in which the subject consistently selects, say, apples over pears regardless of variances in color, apparent ripeness, etc. But things are not so simple. Preferences change over time, which means that one cannot simply assume that tallying up instances of selection will yield a trustworthy generalization about the subject’s
standing preferences. Indeed, the subject might have been expressing nothing more than a preference for each particular apple that was offered at the time it was offered over the particular pear that was offered at the same time.

We can never present a non-linguistic subject with a choice between pearhood and applehood in the abstract in order to determine whether it generally prefers pears or apples. Language helps us to do just that, by enabling us to ask a subject, “Which do you prefer, pears or apples?” So, if one grants that the only way to solve issues concerning holism and indeterminacy is via judging subjects’ preferences by understanding their selections of items at the intensional level, matters are indeed simpler when our subjects are linguistic.

But I think Davidson is too quick to turn to Ramseyan methods to deal with the large number of competing patterns of attribution. I agree that the holism of the mental allows for mixing and matching of beliefs and desires to create many candidate rationalizations of a subject’s behaviors. And I agree that we need a principled way of eliminating many of these potential patterns of attributions. But I do not think the solution is to be found in offering a subject choices between items in order to determine its preferences.

I have already pointed out that interpreters should limit themselves to attributions that contain concepts that a subject could reasonably be expected to mobilize. For Burgean reasons, linguistic subjects can mobilize concepts that they do not fully grasp. But for non-linguistic subjects, the ability to mobilize a concept involves the subject having opportunities to interact with items that fall under the concept, and the fact that
various items did or did not fall under the concept mattering in some way to the being’s interests. For this reason, the number of potential patterns of beliefs and desires available to rationalize a non-linguistic subject’s activities will be significantly fewer than that of the potential rationalizations for a linguistic subject’s activities.

Furthermore, the preferences of non-linguistic beings are likely to be less complex than those of linguistic beings. The fact that the non-linguistic being has a more limited conceptual repertoire also limits the extent to which it will view an object in multiple ways, some of which present the object as desirable, some not. The interests of a non-linguistic being are presumably simpler and more directly tied to survival than those of a linguistic being. They will prefer objects on the basis of the features that matter most for their humble purposes. This gives us a reason to interpret them as mobilizing concepts that capture those features and not others.

Ultimately, I think that the intensionality-related interpretive problems that are supposed to be solved by appeal to language use in the Holism Argument are actually exacerbated by language use in the first place. So, while we cannot appeal to linguistic means to solve these problems when our subject is non-linguistic, the problems are also easier to solve when our subject is non-linguistic. I think we need to appeal to simplicity in order to sort through the various interpretive options for linguistic subjects; for non-linguistic subjects, only the simplest interpretations are on the table in the first place.

Consider again our two imagined subjects, one of whom wishes to see the Morning Star while avoiding seeing the Evening Star, the other of whom wishes to see the Morning Star while avoiding seeing the Morning Star. We needed to capture the
difference between the manners in which these two think of Venus in order to explain why only one of them is clearly irrational. We should choose to attribute to these subjects the simplest ways of representing Venus that make sense of the difference between them.

But with non-linguistic subjects, our needs are simpler in the first place. Non-linguistic subjects are not going to wish to see Venus when conceiving of it one way, but wish to avoid seeing it when conceiving of it another way. The introduction of different linguistic modes of presentation is the source of the difficulty encountered by our astronomically curious subject. The person who wishes to see the Morning Star while avoiding seeing the Evening Star is in this predicament because he has learned two names, “Morning Star” and “Evening Star”, and he does not realize that they name the same object. Language use by the subject allows interpreters to make more complex discriminations, but in many cases it also forces them to need to make more complex discriminations in the first place.

Consider how matters go with a non-linguistic subject. A wildebeest is the same thing as a gnu. Suppose a pride of lions have been stalking some wildebeest. We want to attribute to one of the lions a desire that the lions catch a wildebeest. Under what circumstances will it become important to ask whether the lions conceive of the wildebeest as wildebeest or as gnus? That kind of fine discrimination could be important with a linguistic subject, who might know the words “wildebeest” and “gnu” but not realize they name the same animals. It is not going to be an issue with the lions.
Nor is it going to be an issue for an attribution to Davidson’s canine character of a belief that the pack leader is home. Suppose for the sake of argument that there are three alternative ways of capturing the dog’s manner of representing its master. Which of those options we choose to use when we attribute a belief to the dog will not be important unless the dog can behave in ways that would force us to revise or reconsider an attribution given using any of them. A non-linguistic subject typically is not going to be able to behave in ways that force us from the simplest interpretations. Our standards of subtlety for concept attribution can be relaxed accordingly. We already do this in cases involving linguistic subjects when the behaviors in question do not demand a more sophisticated approach. For any subject, the default interpretation is the least sophisticated one from among the class that renders the being’s activities clearly explicable in rational terms. It is only when a subject’s pattern of moves requires us to discriminate more finely that we need to do so.

But suppose the dog does behave in ways complex enough that an interpreter feels pressure to choose one of the simple interpretations over the others. In that case, the dog will also have given the interpreter a way to distinguish one of these attributions as more appropriate than the others. If a non-linguistic being has made a pattern of moves complex enough to require an insistence on a specific interpretation, that also means that it has made a pattern of moves complex enough for an interpreter to produce and justify a specific interpretation. The degree to which a fine-grained interpretation is possible will necessarily correlate with the degree to which a fine-grained interpretation is needed.
For these reasons, issues of intensionality are not as troubling when it comes to non-linguistic subjects as Davidson believes.

*Intensionality and the Generality Argument*

These considerations do not yet fully address the worry raised by the Generality Argument, however. The Generality Argument is structured as follows.

1. The specific conceptual composition of a thought is determined by its place in a network of beliefs.
2. The network must contain general beliefs if it is to identify the specific conceptual composition of a thought.
3. Only when a subject uses categorical terms do we have any way to make sense of the distinction between general beliefs and conjunctions of beliefs about particulars.
4. So, only linguistic subjects (i.e., subjects who use categorical terms) can have thoughts with determinate conceptual contents.

The intensionality requirement is captured by line 1. And I have no major complaints about the way in which intensionality is invoked here, though again I maintain that connections to perception, action, and linguistic community can also impact thought content. The premise at line 2 contains a worry not raised by the first two arguments. It presents us with a demand not only for a general network of beliefs, but also for a network of general beliefs. The premise at line 3 holds that no sense can be made of the distinction between true generality and mere accumulated particularity apart from linguistic terms that distinguish between them.
To understand my response, it will be helpful to reflect on the ways in which generality is present in any thought. I have set up the Generality Argument to reflect Davidson’s skepticism about our ability to identify logical structure in the cognitive activity of non-linguistic beings. Of course, his point is not just that we cannot work out a logical structure that is present but beyond our ken, but that there is simply no reason to think that logical structure is present in those beings’ cognitive processes. I grant that he has identified a reason for concern. Suppose there are exactly ten trees, and I believe of each one of them that it is a living thing. Do I also believe that trees in general are living things? This question asks about my attitude toward trees as a category, and not just about my attitude toward the individual members of the category. Language allows us to label the relevant category with a word, “tree”, so that we have a way of distinguishing between attitudes taken toward trees in general and attitudes taken toward every individual thing that is in fact a tree. In the absence of language, one might worry that no sense can be made of this distinction, and hence that there is no logical structure there to be found.

The worry is not just that certain kinds of complex thoughts would be unavailable to a non-linguistic being. Even simple thoughts that attribute a property to a particular object are threatened, because their contents contain a categorical concept. If I judge of an item before me that it is a dagger, I have applied a general concept, DAGGER, to the item. So, even simple thoughts about particulars bear contents that require us to make sense of the thinker as trafficking in generality. And Davidson complains that we have

———
134 See Evans (1982).
no way to make sense of a non-linguistic being as doing so. If that is right, then we cannot make sense of the subject as applying concepts, and so we cannot make sense of the subject’s cognitive activities as meeting the intensionality requirement.

The difference between a truly general belief and a conjunction of particular beliefs can show itself in the way new instances are handled. Consider again the example in which there are exactly ten trees, and I believe of each of them that it is a living thing. Suppose my horizons expand, and I discover additional trees. If I merely believed a conjunction of particular claims all along, I will react differently to the new trees than if I believed a general claim all along. If I believed that trees in general were living things, I would proceed under the presumption that these new trees were also living things. On the other hand, if I believed merely that item A was both a tree and a living thing, item B was both a tree and a living thing, etc. then I would not be in a position to presume that the new trees were living things. At most, I might suspect as much given that I had never encountered a tree that was not a living thing. So, we can make sense of an appreciation of generality in the cognitive activities of a being as long as we can make sense of the being as presuming that a new token of a type will have some characteristic of items of that type.

Donald Weiss’s non-linguistic character Arthur is an instructive example here.135 Recall that Arthur has frequently been observed placing metal items in his fire, then hammering them into different shapes. But one day, he finds a new metal item that resists his hammer’s blows after the usual amount of time in the fire. According to

135 For more details, see my discussion of Arthur in Chapter 3, in the section “Considerations Davidson Marshals in “Rational Animals””.

156
Weiss, Arthur’s behaviors (pacing, sitting very still, and bearing an agitated countenance) suggest that Arthur was puzzled to find a metal item that did not become malleable when heated in the fire. The fact that Arthur became puzzled shows that he presumed that the new metal would become malleable, and this suggests that Arthur held a general belief that metals become malleable when heated in Arthur’s fire.

Obviously, Weiss’s example cannot be declared a success as long as any of the ten arguments for linguistic priority remain unanswered. But I believe that it can be declared a tentative success. That is, I believe that a being that behaves as Arthur is supposed to would be a good candidate to be a thinker, and that we would need some overriding reason, provided by the success of one of the ten arguments, to declare that such a being does not think. But more must be said in defense of even this tentative verdict.

The *prima facie* success of the Arthur example hinges on a key feature of the case. First, it is important that Arthur’s body moves in ways identical to the actions a puzzled thinker would be expected to make in Arthur’s circumstances. Just as the Davidsonian radical interpreter and the Quinean radical translator must use body language to judge assent even in the lack of knowledge about what has been assented to, so too the Weissian interpreter uses body language to judge Arthur’s moods during his encounter with the new metal. This provides justification (admittedly, defeasible justification) for claiming that Arthur has become befuddled.

As I pointed out in Chapter 3, Davidson has two options with respect to Arthur. He can either hold that Arthur’s behaviors do not warrant attribution of thought, or he can
hold that the behaviors Arthur is supposed to display could only be displayed by a linguistic being. The first verdict would require backing by a successful argument for linguistic priority; Arthur’s behaviors are indistinguishable from the non-linguistic actions a linguistic being would make when applying a generalization, when becoming puzzled, etc. So, we should judge them similarly unless there are overriding reasons not to do so.

The second verdict would require a different type of defense altogether, some way of demonstrating that language use alone could allow a being to display movements that are indistinguishable from the non-linguistic actions of a linguistic thinker. I strongly doubt that such an argument will be forthcoming. But even an argument of this kind would not rehabilitate the Generality Argument. The problem is with the premise at line 3: Only when a subject uses categorical terms do we have any way to make sense of the distinction between general beliefs and conjunctions of beliefs about particulars. Regardless of whether nature would ever create a non-linguistic being that behaves like Arthur, the fact that we can imagine one is already enough to demonstrate that line 3 is false. Line 3 claims that the use of categorical terms provides our only way of making sense of a distinction between general beliefs and conjunctions of beliefs about particulars. That is not true. The kinds of behaviors Weiss describes show that we have another way of making sense of that distinction: by attending to the circumstances under which the differences between those kinds of belief would make a difference to a being’s response to a new instance.

136 Davidson’s Swampman would stand as a counterexample to such an argument. See Davidson (1987).
If non-linguistic beings who behave in ways similar to Arthur turn out to be impossible, that will not show that we have no way of making sense of a non-linguistic being having a general belief instead of a conjunction of particular beliefs. All it would show is that the difference between those kinds of beliefs would be hard to recognize in beings who behave in less suggestive ways than Arthur is supposed to. That would be an epistemic problem. Davidson is attempting to draw a metaphysical conclusion. He can only complete the inference if the epistemic failure ruins our ability to even make sense of there being a metaphysical determiner present. An interpretationist should not insist that metaphysical contributors to determination of thought exist only when they are practically detectable by an interpreter. Even an interpretationist inclined toward behaviorism would only insist that semantically relevant facts must be detectable in the subject’s behavior by a radical interpreter in principle.

To put it another way, even if metaphysical determination of thought is exhausted by discernment of determiners under conditions that idealize away from epistemic limitations, that only licenses an inference from lack of discernment under non-idealized conditions to lack of determination in cases where the failure to discern ruins our ability to make sense of the notion that there would be determiners under idealized conditions. But in the case of general beliefs, we do know what to look for: behaviors that indicate a presumption that a new instance will be like old ones. With full information about a creature’s behaviors, we should be able to determine when its expectations are met and when they are not. That claim would be defeated by an independent argument against the claim that a non-linguistic being could make presumptions. But it would need to be
defeated by such an argument, and short of the success of one of the ten arguments for linguistic priority, I do not see how this would happen.

Thus, the Generality Argument fails at line 3. The intensionality requirement is legitimate, but it can be met by the cognitive activities of non-linguistic beings. The intensionality-related fact that attributing thought contents composed of specific concepts involves holding the subject accountable to the standards for those concepts’ application still looms as a challenge to my view. But this issue presents challenges that go beyond those of the standard intensionality worries, and I will deal with it last.

The Extensionality Requirement

The intensionality requirement also implies an extensionality requirement. If we must determine how a being thinks of an object in order to determine that it is a thinker, it follows that we must be able to determine which object the being is thinking of in the first place. The extensionality requirement figures most prominently in the Brandomian Syntactic Priority Argument. But Davidson discusses the identification of objects of thought as one of the problems triangulation is intended to solve, as well. My response to Brandom’s argument involves adapting the resources of Davidsonian triangulation to cover cases involving a lone subject.

The Brandomian Syntactic Priority Argument is as follows.
1 Thoughts must bear conceptual contents that represent a world of objects as having certain properties (i.e., to be a thinker, one must confront a world of propertied objects).
2 Only a being who participates in a linguistic practice with subject-predicate structure confronts a world of propertied objects.
3 So, only a being who participates in a linguistic practice with subject-predicate structure is capable of thought.

The premise at line 1 captures both the intensionality and extensionality requirements. I have no complaints against the application of these requirements in line 1. Line 2 insists that the requirements can be met in only one way: via the subject’s participation in a linguistic practice. The guiding idea is that the object-property structure we find in the world is an artifact of the subject-predicate structure of language, which is itself abstracted from the norms governing speech acts.

I am not completely unsympathetic to the idea that the structure of language could play a role in determining the structure of the language user’s world. Given the appeal I made to the notions of pattern and noise in my argument for interpretationism in Chapter 2, this should not be surprising. For someone with my view, the world does not come with pre-established boundaries that separate chunks of reality’s spatiotemporal spread into distinct objects. There is order to be found, but no one particular way of drawing boundaries can claim to be the way the world really is. So, one possibility would be that the manner in which a language user draws boundaries derives from, or at least is conditioned by, her induction into a linguistic practice. And that seems reasonable. Because language has a representational dimension, when one uses a language, one treats the language as capable of representing the world. So, if the norms of linguistic practice
push one to interpret the language as having subject-predicate structure, then a user of the language will be committed to treating the world as having object-property structure.

So, my complaint against Brandom is not that the object-property structure of the language user’s world could not be determined by the subject-predicate structure of the language. My complaint is with the claim that only the subject-predicate structure of a language could bestow object-property structure upon the world. I believe that Davidson’s simple, non-linguistic triangulation suffices to locate an object as a tentative object of thought (pending demonstration that the subject meets all other requirements for being a thinker), and it does not involve the use of subject terms. I have already defended the claim that we can make sense of intensionality for non-linguistic subjects. So, given a method for determining an object of thought, we are in position to make sense of certain properties of the object as the ones in terms of which the subject thinks of the object. Thus, we can handle both the extensionality and intensionality requirements without recourse to the subject-predicate structure of language.

As I discussed in Chapter 3, identification of an object of thought is one of the applications of Davidsonian triangulation. Though Davidson holds that fully linguistic triangulation is required for attribution of thought, linguistic resources are not required in order to carry out the aspect of the attribution that deals with selection of the object of thought. Therefore, Davidson’s simpler, non-linguistic version of triangulation will suffice for our purposes.

Recall that Davidson’s method for identifying an object of belief appeals to causation. Davidson holds that the content, and hence the object, of a simple perceptual
belief is derived from the belief’s cause. But the notion of a belief’s cause is too loose to specify a unique object. According to Davidson, nothing warrants a choice of any particular aspect of the total causal chain that leads up to a lone creature’s responses as the cause of those responses. With two creatures present, however, when each responds not only to a worldly causal chain but also to each other’s responses, an interpreter can trace out the Davidsonian triangle from the two responding creatures to the point at which the causal chains leading to the two creatures’ responses intersect. This point of intersection singles out one public item in a shared world as the aspect of the total cause of each creature’s responses as the cause of both creatures’ responses.

Because this form of triangulation is social but non-linguistic, I could simply appeal to it in order to show that the extensionality requirement can be met by subjects who do not use subject terms. I do think that Davidson’s basic triangulation suffices to justify rejection of line 2 of the Brandomian Syntactic Priority Argument. But I also think that even basic Davidsonian triangulation places unnecessary and insufficiently motivated restrictions on good interpretive practice. I do not think that good interpretation should be limited only to social subjects, and I think that one can capture all the extensionality-relevant features of triangulation within the cognitive activities of a lone subject. So, I will sketch an account of object-directedness that adapts basic Davidsonian triangulation for application to a lone subject.

The key feature of basic Davidsonian triangulation that allows it to handle extensionality concerns is the divergence in perspective between the two beings. If we have only one causal chain leading to one set of responses, the looseness of the notion of
cause prevents us from identifying a unique object as the object of thought. The presence of two causal chains, each leading from some shared source to two different sets of responses, allows for completion of the triangle, and hence identification of the object of thought at the shared source. Therefore, if we can have divergent causal chains within one subject, this will suffice to handle extensionality worries.

I believe that multiple sensory modalities provide the requisite divergence in causal chains that is needed to make sense of an intrasubjective version of triangulation. In cases of intrasubjective triangulation, a single subject responds to information obtained from multiple sensory modalities. The sensory modalities respond to the event in the world, and the subject is prompted to behave by their responses.\textsuperscript{137} Suppose a subject has two sensory modalities. Each sense has the function of causing the subject to behave in ways that are appropriate for survival in its environment. But the two senses react to different features of objects; one processes reflected light, the other processes sound waves. If a bell is rung in front of this being, the being’s response will be the result of at least two distinct causal chains, one leading from the ringing bell through the visual system by means of reflected light, the other leading from the ringing bell through the auditory system by means of sound waves.

Alter the example so that one sense is responding to a stimulus in ways that typically trigger behaviors appropriate for most situations where proposition \( p \) holds. The other sense is responding in a way that typically triggers behaviors appropriate for

\textsuperscript{137} Technically, this would form more of a kite-shape than a triangle when diagrammed, but as the main point is just to adapt Davidsonian triangulation in a way that still preserves its basic uses, I will continue to call it intrasubjective “triangulation”.
most situations where ~p holds. In a scenario like this, we have a divergence between a causal chain that urges the creature toward one set of behaviors and a causal chain that encourages another set of behaviors.

An interpreter might have ways of detecting that the subject is being torn between prompts to behave in incompatible ways. The interpreter might notice that conditions are likely to produce a misleading response from a certain kind of sensory organ. This would give the interpreter reason to interpret any apparent confusion or delay in behavior as indicative of the divergence between the prompts of the subject’s senses. The divergent prompts could then be treated as two of the corners of the triangle, allowing the causal influences that led to each to be traced back to an item in the world, just as in Davidson’s intersubjective version of triangulation.

It is important to notice that I have not yet claimed that intrasubjective triangulation can handle all the issues concerning thought to which Davidson applies his versions of triangulation. I have not yet addressed Davidson’s concern that the solipsist’s world can be any size, which motivates a demand that we have some way of showing that the object we are finding in an objective reality is being located in an objective space by the subject, as well. Davidson appeals to triangulation in order to meet this demand; he explains the origins of a subject’s understanding of the distinction between objective truth and subjective belief by focusing on certain kinds of triangular scenarios. Specifically, Davidson focuses on scenarios in which two beings have interacted enough that, in effect, each has come to treat the other’s responses as signs of certain kinds of worldly
phenomena. Grasp of the subjective-objective contrast is supposed to depend on
situations where the usual correlation between sign and phenomenon is broken.

I have not yet attempted to argue that intrasubjective triangulation can replace
intersubjective triangulation for this purpose. In order for intrasubjective triangulation to
handle this task, one might hold that the subject would have to be able to pay attention to
its various sensory modalities themselves, and not just to the information they deliver. I
will explore this idea presently as part of my discussion of the objectivity requirement.
For now, I want to be clear that I have argued only that intrasubjective triangulation can
successfully reproduce the feature of intersubjective triangulation---multiple causal
chains that lead to responses---that is needed in order to tentatively identify an object of
thought. This allows the extensionality requirement to be met in a way that does not
require even the sociality of the basic Davidsonian triangle.

The Objectivity Requirement

We have seen that a subject must meet the extensionality requirement if it is to
meet the intensionality requirement. That is, if there is going to be a way to identify how
a subject thinks about an object, then that implies that there is a way to identify which
object the subject is thinking of. But the extensionality requirement in turn demands a
solution to the objectivity requirement. If there is going to be a way to identify which
object a subject is thinking of, then that implies that there is a way to determine that the
subject thinks of items in an objective space in the first place.
The objectivity requirement demands that a subject have, or be in, cognitive states that have objectively true or false contents. This has at least two important consequences. First, it implies that the cognitive states represent an objective world. (Even Brandom, who insists on a semantics based on inferential relations rather than representational relations, grants that content has a representational aspect.138) Secondly, it implies that the cognitive states, or the being that has them, are bound by the standard of objective truth.

These consequences suggest two strategies for completing an argument against non-linguistic thinkers, each of which involves pairing the Objectivity Requirement with another of the seven requirements. The fact that thoughts must represent an objective world can motivate a demand that a thinker recognize the objective world as such, and a fortiori that the subject must grasp the distinction between subjective and objective. Thus, the objectivity requirement can be paired with the reflection requirement, which demands that a thinker must have the ability to conceive of its own thoughts as such. Additionally, the fact that thoughts are bound by the standard of objective truth motivates arguments based on the idea that normative binding requires institution within a practice. Thus, the objectivity requirement can be paired with the institution requirement. Several arguments against non-linguistic thinkers attempt to exploit the dual connection of objectivity to reflection and institution. These include the Davidsonian Error, Concept Cluster, and Master Arguments. The case I must ultimately make is that a non-linguistic

---

being can have states that represent an objective world, and thus are bound by the standard of objective truth.

As a preliminary, I want to point out that there is no direct inference from “Subject S has cognitive states that represent an objective world and are bound by the norm of truth” to “Subject S recognizes that its cognitive states represent an objective world and are bound by the norm of truth”. If such an inference is to be made, it will only be done by appeal to auxiliary claims about the nature of normative binding, the nature of representation, or the relationship between these aspects of thought. In the subsection “Objectivity and the Reflection Requirement”, I will argue that an interpreter can make sense of a being as having subjective states that represent an objective world despite the subject’s lack of reflective awareness of the subjective nature of belief and its contrast to objective truth. The most difficult motivation to counter is the notion that a subject cannot be accountable to a norm unless it in some way acknowledges the norm as binding. This motivation for reflection develops naturally into a demand for institution: short of explicit linguistic avowal of governance by a norm, only participation in a practice that institutes the norm can count as acknowledgment of the norm. I will reply to arguments that draw inspiration from this source in the subsection “The Institution Requirement and the Standard of Objective Truth”.

Objectivity and the Reflection Requirement

Reflection and the Nature of Practical Reasoning

168
The Davidsonian Error, Concept Cluster, and Master Arguments each attempt to exploit the idea that thinkers must be capable of reflective awareness. This general idea plays out slightly differently in each of these arguments, but in all of them the basic notion is the same: thinkers must be able to conceive of their thoughts as such. In response, my strategy will be to first warn the reader against a certain type of error that is all too easy to slip into when considering these ideas. It is tempting to think that deliberate action must be based on reflective practical reasoning. I will argue that deliberate action can occur as the result of a non-reflective reasoning process, and that the control requirement can be met simply by insisting on a requirement of non-reflective deliberate action. I will show how these considerations disarm the Davidsonian arguments, and I will examine the extent to which a Davidsonian might be able to patch up any of the arguments.

The easy mistake I want to warn against is to treat deliberate action as if it inherently involves conceiving of one’s thoughts as such. One can slide into the mistake in the following way. Deliberate action is the result of, or the final step in, practical reasoning. Practical reasoning involves the formation of an intention to perform an action, based on consideration of what one desires and what one believes. For example, if I desire that I drink some water, and I believe that the best source of water can be accessed by opening the refrigerator door, then if all goes well I will form an intention to open the refrigerator door. When considering this case, there is a temptation to think that my reasoning must have been reflective, that what happened should be described as
follows: I considered something that I thought of as my desire that I drink some water in conjunction with something that I thought of as my belief that the best source of water could be accessed by opening the refrigerator door; I noticed the inferential relationship between my desire and my belief; and I reasoned that the best way to satisfy my desire would be to open the refrigerator door. This way of representing practical reasoning treats it as if it involves the reasoner’s explicit awareness of the logical structure of the reasoning, and the reasoner’s understanding of how beliefs and desires can work together as reasons to form an intention to act.

But practical reasoning is easily represented in ways that do not presume reflection by the reasoner on its own mental states and the logical structure of the reasoning process. Practical reasoners do not need to “engage in reflection on what outcomes would satisfy their desires, together with the likelihood that the different courses of action have of generating those outcomes.” Practical reasoning does not require thinking to oneself along the following lines: “I have a desire that p, and I have a belief that if I perform action A, then p will be the case, so I will form an intention that I will do A.” Instead, it can be accomplished simply by holding an attitude of desire toward proposition p, and an attitude of belief toward a proposition q that specifies that p will be the case if one A’s, where A is an action one might perform. As Peter Carruthers puts it:

One basic kind of practical reasoning syllogism can be represented thus: DES [P], BEL [if I do Q then P], so INTEND [I do Q]. Here the operators ‘DES’, ‘BEL’ and ‘INTEND’ are meant to designate the first-order

attitudes of desire, belief and intention that figure in the episode of practical reasoning. They are not supposed to be second-order representations of these states figuring within the content of the agent’s reasoning process. The only contents that so figure are $P$, $I do Q$, and the conditional, if $I do Q$ then $P$.¹⁴⁰

Desire, belief, etc. are supposed to be attitudes one can take toward propositions. When the attitude one takes toward $p$ is desire that $p$, that is enough to provide some motivation toward bringing it about that $p$. Typically, there is no additional need to recognize that one has taken an attitude toward $p$, and that the attitude one has taken is of desire. Perhaps there can be cases in which a subject has been pressured to repress certain desires for so long that, though the desires remain, the natural push to satisfy them has become so thoroughly inhibited that the subject would actually need to come to an explicit recognition of the desires in order to be prompted to act on them. But in most situations, the fact that the attitude held toward a proposition is one of desire is already a prod to behave, whether that fact is consciously recognized or not. As Carruthers points out, practical reasoning involves noticing relationships between the contents of propositions to which we in fact take attitudes such as belief and desire. It need not further involve explicit identification of the nature of the attitudes we are in fact taking to those propositions.

So, it does not follow from the nature of practical reasoning that practical reasoning, and hence deliberate action, must be reflective. Davidson acknowledges this

¹⁴⁰ Carruthers (2003). Note that Carruthers’ use of “Q” in his representation of the practical reasoning syllogism is a bit odd, since it is not in the proper grammatical place to function as a variable ranging over propositions, yet the “Q” label seems to suggest a sequential relationship with $P$, which does range over propositions. As my description of practical reasoning demonstrates, the oddity is easily avoided and is not indicative of any major flaw in Carruthers’ proposal.
at times, insisting only that a being’s ability to reason non-reflectively depends on the being’s ability to reason reflectively. But this acknowledgment suffices to disarm several of the Davidsonian arguments against non-linguistic thinkers. Consider version 1 of the Master Argument’s first subargument:

Davidsonian Master Argument Subargument 1, version 1:

1 To have a belief that one could readily discover to be false entails that one is potentially subject to surprise.
2 One cannot be a thinker without having some beliefs that one could readily discover to be false.
3 Surprise requires an awareness of a contrast between something that was believed and something one has come to believe.
4 Awareness of a contrast between something that was believed and something one has come to believe requires grasp of the concept of belief.
5 So, to be a thinker, one must have the concept of belief.

The premises at lines 3 and 4 are crucially ambiguous between the Carruthers-style reading on which line 3 is plausible and the reading Davidson needs in order to motivate line 4. Consider first the reading Davidson needs in order for line 4 to be motivated. In order for us to grant that awareness of a contrast between something that was believed and something one has come to believe requires grasp of the concept of belief, we would have to understand the phrase “something that was believed and something one has come to believe” as implying reflection. The phrase has to be read as indicating an awareness of a contrast not only between the proposition to which one once took an attitude of

---

141 Davidson (1975).
belief and the *proposition* to which one now takes an attitude of belief, but an awareness that the *attitude* one took toward those propositions was *one of belief*.

But this reading renders 3 implausible. Surprise certainly involves coming to be aware that the world is one way when one had previously believed it to be another way. And one would not be surprised to find that p unless one had both previously believed something that contradicts p and had the ability to recognize the incompatibility between p and the conflicting proposition. But if I believe at one moment that it’s safe to go back in the water, and at the next moment find a shark bearing down on me, my surprise at finding that it was not safe to go back in the water is going to depend simply on the proposition IT’S SAFE TO GO BACK IN THE WATER in fact having functioned as a belief for me until just now, and not additionally on my understanding that I have certain states that represent an objective reality, and that one of them was misrepresenting it. Surprise involves having certain standing beliefs about the world, and having a revision to them forced upon one suddenly and jarringly. To be surprised, one would have to suddenly become aware that the world was different from the way one in fact believed it to be. But this involves comparing propositions to which attitudes are in fact taken, and not comparing attitudes as such.

The lesson is that we must always be careful to distinguish between demands that a being consider the propositions that in fact serve as the contents of the being’s various propositional attitudes and demands that the being consider the propositional attitudes themselves as such. As long as I take an attitude of belief toward a proposition, that proposition is available for use in various forms of reasoning. I do not have to pay
attention to the fact that the attitude I am taking toward a proposition is one of belief in order for that propositional attitude to play the distinctive role of a belief in my reasoning.

Similar considerations apply to Davidson’s idea that a thinker must understand that beliefs are the kind of things that can be true or false. Consider the manner in which beliefs operate in the cognitive processes of the typical thinker. Under most circumstances, when a thinker takes an attitude of belief to a proposition, the belief does not function any differently in the thinker’s reasoning than it would if it were known to be true. In most cases, belief revision is a largely unconscious process. It does not require a being to actively consider the fact that the belief has failed to meet the objective standard of truth. Occasionally, we stop to consider whether our beliefs might not be true, or the degree of confidence we have in them, but this is not essential to using beliefs to reason practically. The full Ramseyean apparatus to which Davidson appeals is an important tool for the interpreter who seeks to give an account of an episode of complex practical reasoning, but explicit appeal to it is rarely made in the course of reasoning itself.

Reflection, then, does not seem to make an essential contribution to the role of belief, and the ability of a cognitive state to fulfill that role. As a consequence, Davidson’s Concept Cluster Argument fails at line 3:

1 The concepts BELIEF, TRUTH, and ERROR are acquired only as a group.
2 That group of concepts is acquired only in the process of linguistic interpretation.
3 Grasp of the concepts BELIEF, TRUTH, and ERROR is required in order for a being to be in states that function as beliefs.
4 A being must have beliefs in order to have any thoughts.
5 So, only linguistic interpreters can have thoughts.

And Davidson’s Error Argument Schema---regardless of how it is developed into an argument---fails at line 5:

1 Belief is only made intelligible in contexts involving error.
2 The contrast between error and correct conduct exists, or is made intelligible, only in a norm-governed situation (in the case of belief, the relevant norm is set by the standard of objective truth).
3 Objective truth exists, or is made intelligible, only in contrast to the held true.
4 The notion of being held true applies only to sentences, or is made intelligible only in the context of its application to an interpretee’s sentences.
5 Only beings to whom belief is intelligible can have a belief.
6 A being must have beliefs in order to have any thoughts.
7 So, only linguistic interpreters can have thoughts.

Both the Concept Cluster and Error Arguments fail by appealing to a premise that simply states a demand for reflection. This is not to say that there is nothing of interest in the arguments; on the contrary, the invocation of norms in line 2 of the Error Argument is similar to moves made in arguments that I will discuss under the rubric of the institution requirement. Other Davidsonian arguments offer additional motivation for a reflection requirement, however. And this should not be surprising. Given that Davidson himself will concede that not all thought must be reflective, he must have some further reason for demanding that a being must sometimes be reflective in order to count as a thinker. The two remaining versions of the Master Argument’s first subargument each supply a further motivation for a reflection requirement. One of these, supplied by version 3 of the first
subargument, has to do with the control requirement. The other, supplied by version 2 of the first subargument, has to do with the process of forming general concepts. I will begin my discussion with version 2.

Reflection and Generalization

Version 2 focuses on the phenomenon of reflective surprise.¹⁴²

1 Reflective surprise requires one to have beliefs about the correctness of one’s beliefs.
2 To have beliefs about the correctness of one’s beliefs requires that one grasps the concept of belief.
3 To have a general stock of beliefs, one must be subject to reflective surprise.
4 To be a thinker, one must have a general stock of beliefs.
5 To be a thinker, one must have the concept of belief.

I have argued that surprise need not involve reflection on one’s propositional attitudes as such. But the point of labeling a certain kind of surprise “reflective surprise” is to guarantee, de jure, that it must involve reflection. So, I take the first line to be benign. Line 2 seems to me to be obviously true. In order to believe that a belief of mine is correct or not, I have to be able to consider it in its role as a belief; that requires that I have the concept of belief. I also agree with line 4. Line 3, then, is the target of my rebuttal.

¹⁴² I have renumbered the argument from its appearance in Chapter 3, and cleaned up the premises in lines 1-2. In the present context, I see little benefit to keeping the awkward, context-driven numbering and wording that I used when discussing the argument in Chapter 3.
Line 3 links a creature’s ability to have a general stock of beliefs to that creature having been subject to reflective surprise. But what essential role would reflective surprise play in the generation of a general stock of beliefs? If the point is just supposed to be that a cognitive state cannot play the role of belief for a being unless the being is capable of considering it as such, thus exposing the being to the possibility of experiencing reflective surprise, then there is really nothing novel about the suggestion. The argument would draw from the same motivations I have criticized in the Concept Cluster and Error Arguments. But the reference to the generality of the stock of beliefs is unique to this argument. In order for this argument to go beyond those that I have already rejected, generality must play an important role. And one might suspect that some aspect of the process of generalization requires reflective awareness.

As I have regimented the argument, however, line 3 is not quite what one would really want in order to pursue that suspicion. Line 3 refers to a general stock of beliefs, not a stock of general beliefs. I have represented line 3 this way because it is faithful to Davidson’s phrasing in “Rational Animals”.¹⁴³ But presumably, some liberties can be taken in order to make a friendly revision to the argument. So, take line 3 to claim that the formation of general concepts from experiences of particulars requires the ability to reflect on one’s cognitive states as such.

But even with this alteration, the subargument does not seem very compelling. Presumably, the idea would be that the only way to get to a general concept of, say, HORSE, would be to reflect on all my mental representations of individual horses, and

¹⁴³ See Davidson (1986c).
notice that they all represent the same sort of thing. While such an account might approximate the actual nature of the process by which we form general concepts, it portrays the process as more conscious than it typically is. Perhaps I do come to grasp a concept such as HORSE only because I have been exposed to sensory input from horses (or public representations of horses), and because my brain furnishes me with the ability to sort these inputs as indicative of a certain kind of horse-ish pattern that can be embodied by different particulars. But this will only rarely involve me consciously sifting through my memories of horses in order to try to find a commonality between them.

I argued in Chapter 3 that the chain of thought that leads to this version of the Master Argument is confused. And I believe that this is the reason why this version of the argument does not seem very compelling. It is not an argument that Davidson ever intended to give. But perhaps it can be remedied if one finds a different reason why having a general stock of beliefs would depend on reflective awareness. And maybe the reason is supposed to be that some degree of reflective awareness is required in order for a being to be subject to the rules for the application of the specific concepts that figure in the general stock of beliefs. This is a more interesting suggestion, but it also takes us beyond the reflection requirement to the Kripkensteinian requirement that thinkers must be accountable to standards for concept application. I will discuss that requirement in the section “Institution and the Conceptual Standards Requirement”. For present purposes,

---

144 See the subsection “Arguments for Subconclusion 1”.

178
the lesson is that version 2 of the Master Argument’s first subargument has no independently compelling motivation.

*Reflection and the Control Requirement*

The third version of subargument 1 to the Davidsonian Master Argument is more interesting. This version attempts to motivate the reflection requirement by means of the connection between reflection and the control requirement.

1. Thinkers must be able to take deliberate control of their actions.
2. In order to take deliberate control of their actions, beings must consider how their beliefs advise them to act in the pursuit of their goals.
3. In order to consider how their beliefs advise them, beings must be able to focus their attention on their beliefs.
4. In order to focus attention on one’s beliefs, one must be able to conceive of one’s beliefs as such.
5. Thus, thinkers must be able to conceive of their beliefs as such.

This line of argument does not simply insist that thinkers must be able to reflect on their thoughts. Rather, it aims to identify a specific role in the process of rational action that a thought could not play in the absence of reflection, and argues that a being cannot be a thinker unless its thoughts sometimes play that role. The premise at line 1 insists that thinkers must be able to take deliberate control of their actions. Line 2 claims that beliefs must be treated in an advisory capacity in order for a being to exercise such control. Line 3 points out that one must pay attention to an advisor in order to treat the advisor’s advice as such. This motivates the reflection requirement as it appears in line 4:
the reason a thinker must be reflective is that a thinker must be able to focus attention toward beliefs as states that play an advisory role with respect to action.

My strategy in response will be twofold. I do not think that deliberate control requires reflection. I think that it only requires an ability to consider possible states of affairs and possible behaviors “offline”, and act only once one has had a chance to take attitudes toward these states of affairs and behaviors. But in case my argument is found to be unconvincing, I will also suggest ways in which a non-linguistic creature might come to grasp the contrast between subjective belief and objective truth. I think this kind of response is important, as it also offers a counter to the intuitions that motivate the second subargument to the Master Argument and a slightly reworked Concept Cluster Argument.

I have already discussed the ease with which one can slip into treating practical reasoning as if it must be reflective. It is even easier to suppose that only reflective beings could be capable of exercising deliberate control over actions. Most actions are performed rather automatically and without much conscious thought. But for any behaviors of a being to count as actions, the being must have some capacity for taking control of its own behaviors; otherwise, the move to treat the being as an agent performing actions rather than just as a creature carrying out unthinking behaviors will be under-motivated. To take control of one’s behaviors, one must be able to pause before acting, rather than just act straightaway as a result of cognitive processes to which one is largely inattentive. It is easy to suppose that, during this pause, the deliberate agent must think about what it believes and what it desires, and in effect must ask itself, “Which
course of action makes the most sense given the desires I have and my beliefs about how they may be satisfied?”

As was the case with practical reasoning, I think this way of describing the process invokes reflection where reflection need not be invoked. I find this line of thought compelling insofar as it demands that in order to be an agent who acts, in contrast to a creature that merely behaves, a being must be capable of pausing to consider its options. I balk at the demand that the consideration of options has to be a reflective process.

I have pointed out that practical reasoning can be accomplished by holding an attitude of desire toward proposition p, and an attitude of belief toward a proposition q that specifies that p will be the case if one A’s, where A is an action one might perform. On this model of practical reasoning, the process of deliberation can be understood to demand only that the subject entertain more than one item in the A-role. That is, the deliberator must entertain at least two propositions to which the deliberator takes the attitude of belief, one that specifies a state of affairs that will hold if the being A1’s and one that specifies a state of affairs that will hold if the being A2’s, where A1 and A2 are different actions the deliberator might perform.

The deliberator must consider propositions that are in fact believed and that contain concepts that refer to behaviors that are in fact actions. The deliberator need not have a general understanding of what beliefs and actions are. The deliberator need only take the attitude of belief toward the propositions, and entertain them as part of the process that leads to behavior.
If this is unclear, consider the following two scenarios. In each scenario, a being has a desire that p, a desire that q, a belief that p will be the case if the being A’s, and a belief that q will be the case if the being B’s. We need not suppose that p and q are inconsistent, but it will make it easier to see my point if we suppose that the actions that would bring about the one are different than the actions that would bring about the other.

In scenario 1, the being reflects on her beliefs and desires, decides that satisfying her desire that p is more important than satisfying her desire that q, and thinks to herself, “Well, since A-ing will bring about p and I’ve decided that bringing about p is most important to me, I’m going to A now.” In scenario 2, the being entertains the two relevant propositions to which she holds the attitude of belief. She does not reflect on the relative strengths of her desires, but instead envisions p being the case, and envisions q being the case, and in fact feels more strongly attracted to the former. As a result of the greater strength of her attraction to a state of affairs in which p holds, and her taking the attitude of belief toward the proposition that p will be the case if she A’s, she stops entertaining the beliefs concerning the two courses of action and A’s.

Only in scenario 1 does the subject reflect on her own mental states as such. Yet scenario 2 seems to contain all the elements one could reasonably demand in order to describe the subject as exercising deliberate control over her behavior. She does not act automatically, or purely on the basis of instinct. She entertains propositions that specify different courses of action she might take, and this process leads to the subject behaving in one way rather than the other.
The exercise of deliberate control over actions, then, need not involve reflection. A deliberator must focus attention on her thoughts, and those thoughts must function as reasons for acting one way rather than another, but this process can be carried out without the subject thinking about thoughts, reasons, and actions as such. The deliberator’s focus of attention on her thoughts is a matter of entertaining the contents of those thoughts, and not a matter of recognizing that she has subjective states that she must consult for advice on how to behave. The fact that an attitude toward a proposition p is one of desire, and that an attitude toward a proposition that p will be brought about by A-ing is one of belief, suffice for those attitudes to take on the role of a reason for action when they are entertained jointly. Appropriately related belief-desire pairs are advisory by nature, whether or not their nature is recognized.

At this point, one might object that my account avoids version 3 of the first subargument only by encountering another problem. After all, I am granting that a deliberator must entertain various propositions, and one might complain that the means by which a non-linguistic being would entertain a proposition are mysterious at best. An opponent of mine might here echo Frege’s claim that we only encounter propositions clothed in language.\(^{145}\) If a linguistic medium is required in order to entertain a proposition, then the account of deliberation I have just offered would still render only linguistic beings capable of deliberation.

Perhaps Frege is right that whenever we set ourselves to the (rather bizarre) task of trying to do something akin to observing a proposition, we inevitably conjure up

\(^{145}\) See Frege (1956).
mental imagery akin to seeing a sentence with the mind’s eye, or hearing a sentence spoken in an internal monologue. One might take this to reveal that language is the medium in which we think, or that propositions cannot be entertained except when they are embodied by sentences.

But there are good reasons to reject such diagnoses. The process of thought itself occurs regularly in the absence of any accompaniment from either the sentences of a public language or the imagery of internal monologue. Furthermore, the very notions of an internal monologue and a mind’s eye that can perceive imagined text are no more than gestures in the direction of phenomena we have found no good way to describe. I think the more reasonable conclusion to draw is that the phenomenon of internal monologue, rather than revealing the nature of thought, is merely an artifact of the bizarre attempt to render one’s own thought process observable. In our attempt to force something that is not normally observable, the thought process, into a form in which it is observable, we append to it some quasi-observable mental imagery, and then confusedly portray the thought process as identical to, or embodied by, the imagery.

I do not mean to deny that our thoughts are, in a sense and to a limited extent, introspectible. We do have the ability to portray our thoughts as inner monologue. And the activity in which someone portrays her own thoughts as inner monologue is relevant to proper interpretation of her thoughts, for it reveals how she might honestly portray her thoughts to others. So, when acting as a self-interpreter, each of us has practical access to some interpretation-relevant facts that other interpreters do not. But inaccurate portrayals of thought are possible, even within one’s own inner monologue, as occurs in the
phenomenon of self-deception. The type of introspection of which we are capable does not
ever grant us direct observation of the propositions we entertain, clothed in language
or otherwise. Even public language does not give us observational access to
propositions, but only to sentences that, when all goes well, express them.

The important point I want to stress based on this discussion is that, when I talk
about a subject entertaining a proposition to which she holds the attitude of belief, I am
not envisioning the subject holding up a piece of inner script before the mind’s eye, or
talking her way through a bit of inner monologue. And I firmly reject any attempts to
treat the entertaining of a proposition in deliberate reasoning as if it must somehow
involve something akin to observation of a sentence, perhaps in an inner monologue,
perhaps in a public medium, that expresses the proposition.

I do not claim to have some way of showing that non-linguistic beings entertain
propositions in the same way in their deliberative activities that linguistic beings do in
ours. I do not claim to have an account of how non-linguistic beings might entertain
propositions as part of a deliberative process. I do not even claim to have a clear idea of
how linguistic beings entertain propositions in typical instances of deliberation. But I do
claim that it is not accomplished by anything resembling observation of the proposition.
And I think that the conflation of entertainment of a proposition with observation of a
sentence has to be discouraged if we are to see clearly that deliberation need not involve
linguistically-clothed propositions.

This is not to say that we have no means of judging whether a being has acted
with deliberate control. Online, instinctive, or hard-wired behavioral processes will tend
to eventuate immediately upon receipt of a triggering stimulus. They tend to be inflexible with respect to certain variables that the creature is not wired to handle, and often this inflexibility is immutable. So, an interpreter who seeks signs of offline, controlled elements in a behavioral process may find them in such features as the ability to delay response, and the plasticity to adapt behaviors to new variables. Such features are not as impressive as those Weiss describes in his character Arthur, but whereas there is room for my opponents to question whether a being like Arthur could exist, the features I indicate here can actually be observed in real, non-linguistic beings.

These considerations, I believe, demonstrate that there is no good reason to deny that non-linguistic beings could meet the control requirement. The deliberative process needed for such control does not require reflection, and we can observe non-linguistic beings showing signs that such a process does occur. The main impediment to treating this evidence as such---the conflation of the entertaining of propositions with the observation of sentences---has been shown to involve a confusion.

These considerations also suffice to disarm the Brandomian Deliberation Argument, which contains the premise: Deliberation involves considering what commitments would be attributed by linguistic scorekeepers under various conditions. Properly understood, deliberation does not even involve considering one’s thoughts as such. It certainly does not require one to think about the nature of a linguistic game of reasons.

*Reflection and the Standard of Objective Truth*
So, a reflection requirement cannot be motivated by the nature of practical reasoning, and it cannot be motivated by the need for deliberate control. One might maintain, however, that the real need for reflection is driven by the need to hold a subject accountable to an objective standard, truth. The demand that a subject be able to reflect on its thoughts is not made because of the importance of attending to thoughts as such during the process of reasoning, but because of the importance of contrasting the subjectivity of thoughts with the objectivity of the standard to which the being is held. It would be inappropriate, one might argue, to hold a being accountable to a standard of objective truth when that being has no inkling of such a standard. Most of us would not want to hold a non-linguistic being accountable to moral standards, and our reason for that would be that we doubt that a non-linguistic being would have any way of recognizing such standards. Similar considerations ought to hold for the standard of truth, one might claim.

This way of motivating a reflection requirement promises to rehabilitate some of the Davidsonian lines. The Master Argument, for example, has been defeated in all three versions of its first subargument. But our new motivation for a reflection requirement allows the first subargument, with its focus on the grasp of the concept of belief, to be jettisoned in favor of an expansion of the second subargument that focuses on the need to grasp the subjective-objective contrast itself. And the Concept Cluster Argument can be altered so that grasp of BELIEF, TRUTH, and ERROR are needed not in order to have a
state that functions as a belief, but to be properly held accountable to the standard of truth.

For these reasons, I think it is important for me to point out ways in which a non-linguistic being might come to grasp the notion of an objective standard of truth that one’s states can fail to meet. To do so, I will return to my discussion of intrasubjective triangulation. I think the world can present a non-linguistic being with scenarios that could cause the dawning of the notion of a subjective-objective contrast.

Recall that in cases of intrasubjective triangulation, a subject responds to information obtained from multiple sensory modalities. Each sensory modality responds to an event in the world, and the subject is prompted to behave by the responses of the senses. Consider again a case discussed in the Extensionality Requirement section of this chapter, in which a being with two sensory modalities is being prodded to behave in incompatible ways by the senses. In the Extensionality Requirement section, my discussion of this case focused on the ability of an interpreter to recognize that a subject was being urged by its senses to act in two incompatible ways. I noted then that I had not yet attempted to employ intrasubjective triangulation for the task of showing how a non-linguistic being might arrive at grasp of the subjective-objective contrast. I turn to that task now.

Recall that in Davidson’s non-linguistic version of intersubjective triangulation, two creatures in effect come to treat a certain kind of response by the other as a sign that a certain type of event has occurred. The first inklings of a subjective-objective contrast are supposed to emerge from situations where the usual correlation between sign and
event is broken. In order to assess intrasubjective triangulation, we should ask whether this basic setup can be recreated in the intrasubjective case. The first thing to notice is that a creature does not have to come to treat its sensory responses as signs of certain kinds of events; the senses serve their purpose because, in effect, their responses are automatically treated as signs of kinds of events. Of course, the “in effect” is important here; without it, my claim would be question-begging. But my opponents are in no position to complain about this way of phrasing things. Davidson’s non-linguistic intersubjective triangulation must involve a similar treatment of responses as signs in effect only. Otherwise, it would yield the claim that the notion of objectivity arises as a result of a subject coming to think that a certain response is a sign of a certain event. That result is obviously unacceptable given Davidson’s desire to claim that there can be no thought among beings who do not grasp the notions of objective and subjective.

So, there should be no quarrel with my claim that sensory responses are in effect treated by their subjects as signs that certain events have occurred, though it would be both question-begging and false to assert of many beings that they explicitly treat their sensory responses as signs. And if failure of the usual correlation between response and event can stir a being toward grasp of the subjective-objective contrast in the intersubjective case, it should be able to do so in the intrasubjective case, as well. A creature might associate certain stimuli from various sensory organs with the presence of a food item. Suppose for example that pieces of a certain kind of fruit have a distinctive look, smell, taste, etc. and that a monkey associates each of these sensory responses with the presence of fruit. Suppose the monkey receives a visual stimulus that it associates
with the fruit, but attempts to eat the item prove unsuccessful. Upon closer inspection, the proper visual stimulus remains, but the other senses are responding in ways that indicate that there is no fruit here, just a rock.

In a situation such as this, a fully linguistic adult human would no doubt conclude that the item in question is a rock that looks like fruit. If prompted to reflect further on the perceptual beliefs formed upon looking at the rock, the human subject would conclude that he had misrepresented the rock as fruit. We cannot simply suppose the monkey to be capable of this kind of appraisal of the situation. Nonetheless, given that it does normally follow sensory prompts to behave, the fact that all has not gone smoothly in this case will have to register somehow within the monkey’s cognitive processes if it is to stop attempting to eat the rock.

Granted, this will probably be accomplished without any conscious thought on the monkey’s part. But the situation will have to feel different than one in which all had gone well and the item has turned out to be fruit. There will have to be an urge against trying to eat the item strong enough to overcome the visual prod to eat the item. Though it is unlikely that sapience is involved in such a case, sentience is involved.

It is important to realize that interpretationism about the thinking mind in no way precludes realism about the phenomenal mind. An acknowledgment that there is such a thing as conscious experience in no way commits one to treating it as providing introspective access to all thought contents, to holding that it is scanned by a mind’s eye in order to convert its Given content into infallible knowledge, or anything else that would make an interpretationist as such balk. The exact nature of the relationships
between the phenomenal mind and other related items, such as the cognitive mind and neural processes, is less than clear. But there is a phenomenal aspect to some mental features, such as attention, that are relevant to the reasoning process, and I suspect that this is the true origin of the first inkling of the contrast between an objective world and subjective thoughts.

The world in fact sets a standard that the monkey’s visual system fails to meet when it produces the inappropriate behavioral prod. The cognitive conflict between the visual prod and those from the other senses will be felt by the monkey as a phenomenal dissonance that is not present when the senses prompt behavior in a unified way. The feeling of dissonance alone surely will not suffice to indicate the failure of a subjective state to meet an objective standard, but neither will the observation of a correlation failing to hold between one chunk of nature (a worldly event) and another (another creature’s response to it), as in Davidsonian non-linguistic triangulation. And if Davidson’s non-linguistic triangulator is supposed to have some inkling that one of those chunks of nature is similar to itself, this will only prompt the idea of an intersubjective standard if the triangulator already has some inkling of self as subject. If all we are after is inklings, however, the intrasubjective instance of correlation failure, accompanied by phenomenal dissonance, will do the job at least as well as the Davidsonian interpersonal version.

In fact, the phenomenal mind really ought to indicate the difference between objective world and subjective representation more readily than an external failure of correlation could. Alterations to one’s sensory organs result in distortions of the
representation of the world. But they do not result in alterations of the world itself. If a being has only one sensory modality, this would not be readily apparent. But with additional sensory organs that remain highly functional throughout a perceptual episode, the more apparent will be any distortion by another sensory organ. Closing one’s eyes does not make the items that were so recently visually perceived go away---as is often indicated by the persistent detection of such items through auditory or olfactory senses---though it does detract from one’s ability to represent all of their aspects.

Furthermore, the link between proprioception and action provides a basis for a self vs. other contrast. The proprioceptive sense of bodily position might not amount to much in creatures that only behave in direct response to stimuli. But even the simplest being that is capable of pausing for cognitive processing before carrying out a behavioral response to a stimulus must have some proprioceptive inkling, however inchoate, of the fact that some things---its proprioceptively tracked body parts---are among those that move when it, the creature, moves.

Even an inchoate notion of bodily self provides a basis for attending to one’s own sensory organs as one’s own. And this provides a basis for attending to the sensory organs themselves, rather than just to the perceptions they provide. If a being can pay attention to, say, the status of its eyes as open or closed, the difference the eyes’ status makes to what is perceived, and, by appeal to other senses that do not undergo a status change during the perceptual episode, the lack of difference the eyes’ status makes to what is present, then the being is so positioned that grasp of the contrast between objective world and subjective representation can begin to emerge. And none of this
requires an additional being, though it does perhaps require a degree of behavioral plasticity and more than one sensory modality.

This is as it should be. Davidson complains that the solipsist’s world can be any size, and so it is not a world. Then, he attempts to solve the problem by introducing other subjects. But insofar as the worry was that a lone being would have no basis for treating itself as moving within an objective reality, the addition of other subjects will not help. Their status as additional subjects will not be recognized as such by the being that does not grasp the subjective-objective contrast. Since the solipsist’s world can be any size, it can clearly be big enough to include any other purported subjects as well. That is, after all, how one motivates a slide from Berkeley’s idealism to solipsism in a discussion of skepticism about the external world.

The Institution Requirement and the Standard of Objective Truth

The argument I have just concluded embodies a radical departure from the Davidsonian view of the origins of grasp of the subjective-objective contrast. Davidson thinks that sociality plays an essential role in initiating the process of coming to grasp the concepts of subjective thought and objective reality. I have argued that sociality need not, and could not, play such a role. But notice that the Davidsonian has not been defeated yet. For Davidson himself is quite clear that the non-linguistic form of social triangulation is not supposed to suffice for grasp of the subjective-objective contrast, or of any other concepts. It is only when the interactions between the beings take on the
character of linguistic interpretation that conceptualization occurs at all. And the most compelling motivation for this claim is one I have not yet addressed: the claims that normative standards must be instituted in a practice, that only practitioners are properly held accountable to the norms a practice institutes, and that linguistic practice institutes the standard of objective truth. In order to fully demonstrate that a non-linguistic being could meet the objectivity requirement, I will have to show not only that a non-linguistic being could have some inkling of a subjective-objective distinction, but that such a being could be properly held accountable to the standard of objective truth despite its failure to participate in the institution of a public linguistic practice.

In addition to the possibility that this institutional line will rehabilitate some of the Davidsonian arguments, the Basic Institutional Argument embodies the dual concern with objectivity and institution, as well. The concern for objectivity is displayed in the argument’s first premise.

1. Thinkers’ behavioral and cognitive processes must be accountable to standards of truth, rationality, and concept application, not merely describable in terms of such standards.
2. Only beings who have been initiated into a practice are accountable to that practice’s standards (in this case, standards of truth, rationality, and concept application).
3. Only linguistic beings have been initiated into a practice governed by standards of truth, rationality, and concept application.
4. So, only linguistic beings are thinkers.

Notice that the first premise extends beyond the objectivity requirement’s concern with the standard of truth to standards of rationality and concept application, as well. The
additional premises connect the various forms of normative binding with the institution of
the relevant norms in linguistic practice.

As I noted in Chapter 3,\textsuperscript{146} the mere fact that we can describe a being’s behaviors
in terms of various rules is insufficient to justify the claim that the being is accountable to
the standards we have imposed. The fact that a subject’s activities are rule-describable
does not entail that they are rule-accountable. Explicit avowals by a subject concerning
the nature of rules the subject follows suffice to warrant treating the subject as bound by
those rules. Non-linguistic beings are obviously unable to offer explicit avowals as to
what standards, if any, govern their behaviors. In the absence of explicit avowals, we are
often able to make sense of a being as bound by a standard because it participates in a
practice in which the standard is instituted. But the obvious practice in which norms of
truth, rationality, and concept application are instituted is linguistic practice, so this
strategy is unavailable for non-linguistic creatures. Thus, if one is to make sense of non-
linguistic creatures as genuinely accountable to the standards of truth, rationality, and
concept application, another strategy will be needed.

My strategy will make use of four primary observations. One of these is that
survival and flourishing are, by their nature, default ends for any creature, and they
cannot be maintained unless the creature is capable of surrounding itself with the proper
circumstances.\textsuperscript{147} The second is that the likelihood of maintaining the ends of survival
and flourishing typically increases as the creature’s behavioral and cognitive processes

\textsuperscript{146} See the section, “The Institution of Norms in the Social Practice of Language.”

\textsuperscript{147} Notice that these are default ends for the individual creature as such, and not necessarily for the creature
as a member of its species or a transmitter of its genes.
conform to the standards set by truth and rationality. The third observation is that creatures that have the capacity to learn undergo behavioral and cognitive modifications that occur specifically because they allow the creatures to conform more closely to the standards of truth and rationality. The final observation is that the ability to track some categories will correlate more closely to a creature’s survival and flourishing than will the ability to track others, giving an interpreter a reason to identify a subject’s cognitive processes as aimed at representing items as falling under those concepts that pick out the flourishing-relevant categories. I will discuss the first three points in the remainder of this section. Discussion of the final point will be postponed until the section that follows this one.

My first observation is that survival and flourishing are, by their nature, default ends for any creature, and they cannot be maintained unless the creature is capable of surrounding itself with the proper circumstances. I will elaborate on this claim, and my reasons for believing it to be true, presently. I intend to appeal to it as follows. Because survival and flourishing can be treated as default ends for any creature, and they cannot be maintained unless the creature is capable of surrounding itself with the proper circumstances, the creature’s ability to maneuver through the world can be treated on the teleological model employed in rational assessment. Thus, there is a legitimate sense in which the ends of survival and flourishing govern the activities of the creatures, whether the creatures are aware of that or not. But to justify this claim, I need to say more about why survival and flourishing are default ends for any creature.
Clearly, certain conditions must be maintained in order for any creature to survive. Through behaviors such as feeding, drinking, and respiration, it must take in whatever chemical compounds it requires in order to maintain the functions of a living thing. It must have some way of eliminating the harmful byproducts of such processes. And it must avoid being consumed by other creatures. If a creature fails to maintain any of these conditions, it will die.

All else equal, death is a bad thing. There are cases in which a creature will sacrifice itself for the benefit of its offspring, or, in the case of human creatures, for any number of reasons. But these kinds of cases arise only in circumstances that occur rarely in the life of any one creature. The default aspiration of any creature is to continue living, and this aspiration is overridden only in special circumstances. And this default aspiration is not merely to live, but to sustain a fairly healthy existence in which one’s needs are not only being met, but being met well. Flourishing is preferable to merely scraping by. In defense of this claim, I think that little needs to be done other than to point out that flourishing is, by its very nature, a state of being that is inherently valuable to the creature that enjoys it. As is the case with survival, there can be circumstances in which other factors outweigh, or are treated by the subject as though they outweigh, the value of flourishing. But in the absence of special circumstances, the inherent value of flourishing renders it a default aim for any being.

All else equal, then, survival and flourishing are, by their natures, in the interests of any being. Of course, this does not suffice to show that all creatures have an attitude of desire toward the proposition that they live and flourish. But it does provide a basis
for employing a teleological mode of explanation when seeking to understand why creatures behave as they do. This alone does not suffice to show that the behaviors in question are normatively constrained. But it does provide one important aspect of a normative practice. Where there are ends, there can be failures to bring those ends about. The fact that a behavior can fail to bring about a valued end does not necessarily show that the behavior was governed by a rule specifying that the end was to be brought about, but it does provide a basis for an interpreter to assess the behavior in addition to merely describing it.

The second key observation is that the likelihood of maintaining the ends of survival and flourishing typically increases as the creature’s behavioral and cognitive processes conform more closely to the standards set by truth and rationality. I think this is clear enough. True beliefs help a being to successfully navigate the world; false beliefs typically detract from success. And creatures that behave and cognize in ways that suit their preferences, their true beliefs, and the inferential relationships between them will tend to survive and flourish more than creatures whose behaviors and cognitive episodes are irrational.148

I do not mean to claim that there can be no such thing as a useful fiction. Given the finite nature of creatures’ cognitive systems, a complex truth might be overwhelming, whereas a simpler falsehood that approximated that truth might serve the creature’s purposes just fine. And I do not mean to claim that there could never be special

---

148 Thus, the meeting the standards set by rationality involves the four normative constraints on thought that I mentioned in Chapter 1 (see p. 12): consistency of beliefs, transitivity of preference rankings, inferential soundness, and maximization of expected utility.
circumstances in which a certain degree of irrationality might produce better results for a creature than would perfect rationality. My claim is that, in the overwhelming majority of comparisons one might make, a cognitive system that conforms to the standards set by truth and rationality to a relatively high degree will be more beneficial in terms of maintaining survival and flourishing than a cognitive system that conforms to those standards to a lesser degree.

As a consequence, truth and rationality can serve as guiding standards in the quest for survival and flourishing. And this is how they are typically treated by those beings---linguistic humans---who can reflect on, and seek to actively alter, their cognitive systems. At least, the wise ones among us attempt to rid themselves of false beliefs and tendencies to act in ways that do not reflect their beliefs and preferences.

But we cannot assume this level of reflection among non-linguistic beings. We cannot safely assume it of most linguistic humans, for that matter. So, if interpreters are to treat the activities of these beings as accountable to the standards of truth and rationality, there must be something else we can appeal to besides explicit reflective attempts on the part of individual subjects to increase their own conformity to those standards. And here the fact that linguistic humans participate in linguistic practices in which the standards of truth and rationality are applied successfully renders them accountable to those standards.

Because this strategy will not work for non-linguistic creatures, my third observation becomes crucial. The third observation is that creatures that have the
capacity to learn\textsuperscript{149} undergo behavioral and cognitive modifications that are driven by the need to better meet the ends of survival and flourishing. I believe that my first observation (that survival and flourishing are default ends for any creature) licenses us to view the modifications that occur during the learning process from a perspective that focuses on the learner’s interests. Viewed from that perspective, the purpose of the modifications is that they increase the likelihood that the creature will maintain its survival and increase its level of flourishing.

Here, my second observation becomes important: a cognitive system’s ability to maintain its possessor’s survival and flourishing typically increases as the system approaches conformity to the standards of truth and rationality. Cognitive systems can be modified so that they conform better to the standards of truth and rationality. For example, a modification could decrease a tendency to give a false positive or negative in a certain kind of situation, bringing the being’s perceptual system into greater conformity with the standard of truth. Or a modification could decrease a tendency to act in ways that conform to a stronger preference for $p$ than for $q$ across many cases where $q$ is actually preferred to $p$, thus increasing conformity to rational standards.

For my purposes, the key point here is that creatures’ cognitive systems undergo these modifications during the learning process \textit{because} the modifications allow the cognitive systems to conform more closely to the standards of truth and rationality. Those standards govern the process by which the learner’s cognitive system is structured,

\textsuperscript{149} The learning can be associative, and thus does not presuppose thought, as that would obviously be question-begging.
regardless of whether the learner is aware of that fact.\textsuperscript{150} Whether we should call the pattern of cognitive processes and behaviors that result from learning a practice, I do not care to judge. But I do think that a disjunctive conclusion is warranted: the fact that the pattern results from a process that is governed by the standards of truth and rationality justifies us in either counting this kind of pattern as a practice that institutes the standards of truth and rationality, or concluding that a pattern of cognitive and behavioral activities can be properly held accountable to the standards of truth and rationality despite the absence of a practice that institutes those standards.

If we opt for the first disjunct, we should reject line 3 of the Basic Institutional Argument: only linguistic beings have been initiated into a practice governed by standards of truth, rationality, and concept application. According to the first disjunct, going through the learning process counts as such an initiation, and such processes are not restricted to linguistic beings. On the other hand, if we opt for the second disjunct, then we can reject line 2 of the Basic Institutional Argument: only beings who have been initiated into a practice are accountable to that practice’s standards. According to the second disjunct, the fact that the standards of truth and rationality govern the learning process allows us to hold learners accountable to those standards despite the fact that they have not been initiated into any practice that institutes those standards.

I do not mean to claim that the fact that a subject learns suffices to warrant attributions of thought to the subject. I do claim, however, that the fact that a subject

\textsuperscript{150} Driven by selection pressures, a similar process of modification could occur on the phylogenetic level, as well. But because this type of alteration does not merely impact the structure of one individual’s cognitive system, I am not sure that it is appropriate to view a change that comes about in this manner from a perspective that focuses on the individual subject’s interests.
learns renders its activities potentially accountable to the standards of truth and rationality. If we find that the subject meets the other requirements for interpretation as a thinker, we can drop the “potentially”. The fact that the subject is a learner contributes the aspect of thought---accountability to standards of truth and rationality---that most strongly motivates the objectivity and institution requirements for interpretation as a thinker. Thus, I think it answers the arguments against non-linguistic thinkers that are based on those requirements---with one important caveat. The Basic Institution Argument not only mentions the standards of truth and rationality, but also the standards for application of specific concepts. To complete my counter-argument, I will have to extend my strategy to deal with these standards, as well.

Institution and the Conceptual Standards Requirement

Here, I will address an issue that has been deferred twice already in this chapter, the only one of the seven requirements for thought to which I have not yet responded: the accountability of thinkers to standards for the application of specific concepts. I mentioned this requirement both as a motivation for a reflection requirement, and as a reason to doubt that a non-linguistic being could meet the intensionality requirement. It is the last remaining hope for the Basic Institutional Argument. And it provides the primary motivation for the one argument against non-linguistic thinkers to which I have given no reply so far, the Kripkensteinian Argument:

1 Thinkers’ behavioral and cognitive processes must be constrained by standards of concept application.
2 Considered in isolation from linguistic practices, the activities of a subject fail to determine whether the subject is accountable to a standard of concept application that it meets, or to a standard it fails to meet.

3 If it is indeterminate whether a subject’s activities are in accord with the standard of concept application to which it is accountable or in conflict with the standard to which it is accountable, then the subject is not constrained by standards of concept application.

4 If a being was not initiated into linguistic practice, then it must be considered in isolation from linguistic practices.

5 If a being is non-linguistic, then it was not initiated into linguistic practice.

6 So, if a being is non-linguistic, then it is not a thinker.

The requirement of accountability to standards for concept application is stated in line 1. The demand for a being to use a public language in order to meet the requirement is technically made in lines 4/5, but the primary motivation for this demand is embodied by lines 2/3. Linguistic practice is supposed to resolve a particular problem with construing a non-linguistic creature as accountable to standards for concept application, a problem that has to do with indeterminacy. The problem is that indeterminacy threatens our ability to make sense of a distinction between correctly applying one concept and incorrectly applying another. If we have not made sense of a distinction between correct and incorrect concept application, that means that we have not make sense of the notion that the activities in question are accountable to standards for concept application at all.

In response to the worry that non-linguistic beings cannot meet this requirement, I will show how my basic strategy for demonstrating the accountability of non-linguistic beings to standards of truth and rationality can be extended to cover standards for concept application, as well. Again, I will appeal to survival and flourishing as default ends for any creature. The ability to track some categories will correlate more closely to a
creature’s survival and flourishing than will the ability to track others. This gives us a reason to treat the process in which creatures’ cognitive systems are shaped by experience as governed by conceptual standards that demarcate certain categories and not others: those categories that the creature must track if it is to best meet its ends of survival and flourishing. This gives us a reason to identify a subject’s cognitive processes as aimed at representing items as falling under certain concepts.

The (too) rough idea is that we are to identify a subject’s cognitive processes as aimed at representing items as falling under those concepts that demarcate the categories that are most relevant to flourishing. The more refined idea is that we are to begin by identifying a set of concepts the creature might plausibly possess given its perceptual sensitivities and learning environment, and from that set, select the simplest concepts that demarcate the categories that are most relevant to flourishing. These are the concepts we ought to say the being mobilizes, given that it is able to meet the other requirements for counting as a thinker. To the extent that cognitive processes do not properly sort items according to whether they fall under the categories that these concepts demarcate, those processes misrepresent the world.

In the previous section, I argued that the cognitive and behavioral activities of non-linguistic creatures can be properly held accountable to standards of truth and rationality in cases in which those activities were the result of a learning process (and the non-linguistic being meets all other requirements for counting as a thinker). The reason for this is that the learning process serves to modify the activities so that they are brought closer to conformity with the standards of truth and rationality. And the learning process
does this because increases in conformity with those standards increase the likelihood that the creature will continue to satisfy its default ends of survival and flourishing.

Similar considerations apply to the standards of application for concepts. The standards for a concept’s application demarcate a category of items (where “items” is construed as broadly as possible, and not merely as referring to what we would normally call “objects”, but also events, stuffs, functions, etc.) that fall under the concept. Some of these demarcations will divide items according to factors the presence or absence of which are highly relevant to a creature’s survival and flourishing, such as whether or not the items in question are nutritious or dangerous. Other demarcations will divide items in ways that have no particular bearing on whether the creature’s ends are met, such as the division of items into ones that have been within 500 feet of a Volkswagen and those that have not.

Just as modification of a cognitive system so that it more closely conforms to standards of truth and rationality will increase the likelihood that the system’s bearer will continue to survive and flourish, so too will modification of a cognitive system so that it better tracks the categories that are most relevant to flourishing. And just as some creatures’ cognitive systems undergo modifications during a learning process that serve to bring the systems closer to conformity with the standards of truth and rationality, so too do some creatures’ cognitive systems undergo modifications that serve to render the creatures better at discriminating between items that do and do not fall under highly flourishing-relevant categories. And this gives us a reason to treat the process in which a creature’s cognitive system is modified in this way as governed by conceptual standards
that demarcate the highly flourishing-relevant categories, rather than other conceptual standards or none at all. Just as the connection between the default ends of survival and flourishing and a cognitive system’s conformity to the standards of truth and rationality renders the learning process governed by those standards, so too the connection between survival and flourishing and a cognitive system’s ability to track membership in certain categories gives us reason to treat the process in which the creature’s discriminatory skills are honed as governed by the standards for application of concepts that demarcate those categories. This constitutes a reason for an interpreter to treat the creature as accountable to standards for those concepts’ application (again, given that the creature meets the other requirements for interpretation as a thinker).

In response to the Kripkensteinian Argument’s second premise—considered in isolation from linguistic practices, the activities of a subject fail to determine whether the subject is accountable to a standard of concept application that it meets, or to a standard it fails to meet—the rough idea is that we are to identify a subject’s cognitive processes as aimed at representing items as falling under those concepts that demarcate the categories that are most relevant to flourishing. So we are to hold the subject accountable to the standards for those concepts’ application in our attributions; that entails that we will attribute thoughts composed of those concepts and not others. And the hope would be that this restriction of concepts would yield a clear answer as to whether we ought to be attributing a proper application of one concept or an improper application of another one.

151 These are not necessarily the categories that the creature’s cognitive system most closely tracks. This would perhaps be a problem for someone who was out to give a purely descriptive account of concepts in terms of dispositions to respond, (see Fodor (1990) for complaints along these lines), but given the normative/teleological aspects of my account, I do not see it as a problem.
Consider a case analogous to the PLUS and QUUS interpretations of “+”, but altered so that it deals with non-linguistic behaviors instead of use of a linguistic symbol, and so that the topic has a more direct impact on a creature’s survival and flourishing. Consider a subject that has been observed on many indications to engage in a behavior that we would want to describe as “running away from a predator”. Now imagine that on the date January 1, 20XX, we encounter a Kripkensteinian who suggests that the being might be equally well-described as running away from qedators, where “qedator” applies to all and only predators larger than or equal in size to the smallest predator our subject has encountered as of January 1, 20XX (treating “predator” as referring in this context only to those beings who are a predatory threat to creatures like our subject). And suppose that the next day, our subject is stalked by a creature that counts as a predator but not as a qedator. Our subject seems to detect the presence of the other being, but fails to run, and is killed and eaten.

We think that our subject has failed to follow the rule that was embodied in its previous intentions, “run away from predators”, and we speculate that the reason for this failure was that the subject’s cognitive system misrepresented the predator as some other, non-threatening sort of creature, possibly due to its small size. But our Kripkensteinian friend suggests that we have no more basis for this diagnosis than one in which we take the subject’s behavior to accord perfectly well with its intentions all along, which had been to run away from qedators. The account I have proposed rules out attributions involving “qedator”, because discriminating between predators and non-predators is more important to the subject’s survival than discriminating between qedators and non-
qedators. So, the subject’s cognitive and behavioral processes are accountable to the standards set by PREDATOR, but not accountable to the standards set by QEDATOR, despite the fact that the rationalizing explanation appears to go more smoothly when we decide that our subject had a desire only to avoid qedators, not predators, all along.

I think my method, as I have described it so far, works for the predator-qedator case. But it is not sufficiently refined to handle another worry that a Kripkensteinian is sure to have. This worry has to do with the vagueness of the notion that a category is highly relevant to a creature’s survival and flourishing. The vagueness of the notion of a category being highly relevant to a creature’s survival and flourishing creates problems for my account when we move beyond examples that are modeled closely upon Kripke’s plus-quus case. My account is capable of drawing a principled distinction between concepts like PREDATOR and QEDATOR only because the close relationship between those two concepts makes it easy to compare the overall utility of tracking the one category vs. tracking the other. Tracking predators is not just highly relevant to flourishing; it is also clearly more relevant to flourishing than tracking qedators. But a direct comparison is not always going to yield such clear results.

Consider PREDATOR in comparison to PREDATOR-OR-700-ARMED-NINJA, a concept that demarcates the category of things that are either predators or 700-armed ninjas. The behavioral processes that will lead us to want to attribute intentions to run from predators would not be any less successful if they were based on intentions to run from predators or 700-armed ninjas. So, a method of holding beings accountable to conceptual standards that is based strictly on the correlation between tracking things that
fall under those concepts and flourishing will not properly distinguish between a concept C and a concept that includes objects that fall under C and objects that fall under a category that is irrelevant to survival and flourishing. So, my account needs to be supplemented by some other factor that will rule out these cases.

Here, I appeal to simplicity. There is no reason to attribute the concept that takes in 700-armed ninjas as well as predators. Assume for the sake of argument that PREDATOR suffices to take in all the cases that are relevant to the creature’s behavioral success when it comes to the matter of running away. An interpreter would need some motivation to revise away from an attribution containing the simplest concept that suffices to take in the flourishing-relevant cases. The default practice for an interpreter, as with any other explainer, is to start with the simplest account and only complicate it as necessary. If avoiding 700-armed ninjas would be important to the creature’s flourishing, on the off chance that one should appear, then there would be reason to revisit the attribution. If not, then there is no reason to bother with a concept that includes them in our attributions.

There is also another reason to reject an attribution to a non-linguistic subject that contains a concept that takes in predators and 700-armed ninjas. It is doubtful that there would be any way for a non-linguistic being to make much progress toward distinguishing between things that are ninjas and things that are not. To count as a ninja, a person must occupy a specific role in human society; the person must at least be a martial artist and perform mercenary tasks. To sort things into ninjas and non-ninjas,

\[152\] Simplicity here is to be judged relative to the interpreter’s needs as such.
then, requires sensitivity to whether or not a person is trained in the martial arts and whether or not the person performs mercenary tasks. A subject that has no way of even beginning to learn to distinguish between martial artists and non-martial artists, or to learn the difference between a mercenary task and non-mercenary activities, is not going to be able to undergo a learning process governed by the standards for application of NINJA.

This is not to say that a creature must be able to learn to distinguish perfectly between items that fit a category and items that do not in order to undergo a learning process that is governed by the standards for application of the concept. Nor do I mean to imply that a creature must be able to attend to all the factors upon which the demarcation of a category by a concept is based. People had a concept of water before water’s chemical nature was discovered, despite the fact that the category of stuff that counts as water is determined by the chemical nature of that stuff. But I do claim that a creature whose perceptual sensitivities or learning environment are impoverished to the point that it cannot attend to the most basic perceptual cues that indicate the presence or absence of a type of thing, or to the features of the world that metaphysically determine what does and does not fall under that type, is not a candidate to have thoughts containing a concept that demarcates that type.

So, my altered proposal is that an interpreter of a creature is to begin by identifying a set of concepts the creature might plausibly possess given its perceptual sensitivities and learning environment, and, from that set, select the simplest concepts that demarcate the categories that are most relevant to the creature’s flourishing. Given
that the creature meets all other requirements for counting as a thinker, and that it has indeed undergone a process that has honed its ability to discriminate between things that fall under those concepts and those that do not, the creature’s cognitive processes are properly held accountable to the standards for application of those concepts. This fully answers the Kripkensteinian Argument by rejecting its second premise, and shows that the strategy I used to respond to lines 2/3 of the Basic Institutional Argument with respect to truth and rationality can be extended to cover concept application as well. It also eliminates any remaining motivations for the claim that non-linguistic beings cannot meet the intensionality or objectivity requirements, and for the claim that there is a need to meet the reflection requirement.

Conclusion: an Interpretationist Account of the Thinking Mind

In Chapter 3, I identified seven proposed requirements that a being must meet if it is to be properly interpreted as a thinker. I also identified ten arguments against the possibility of non-linguistic thinkers, each of which in some way exploited at least one of the seven purported requirements. In Chapter 4, I have shown that none of the ten arguments actually establishes the impossibility of non-linguistic thinkers.

I have accomplished this within a framework that respects the motivations behind the ten arguments. I have rejected only one of the seven proposed requirements, the reflection requirement, and I have shown how a non-reflective deliberation requirement can perform the legitimate work for which the reflection requirement was intended. I
have left the status of the *institution requirement* unclear because I see room for
disagreement over whether the activities that result from a learning process should be
called a “practice” or not; I have argued that, however that terminological matter is
settled, the learning process is capable of carrying out the legitimate work for which the
institution requirement was intended.

I have accepted the other five requirements---the *intensionality, extensionality,
control, objectivity, and conceptual standards requirements*---and given an account of
how non-linguistic beings might meet them.

The *intensionality requirement* can be met by restricting the choice of concepts
contained in attributions to ones to which the subject has sensitivity and interest. The
relatively low density of a non-linguistic creature’s web of belief is not a problem as long
as the existing nodes preserve the most important links between the reduced pool of
concepts available for attribution. Furthermore, the fact that fewer concepts are available
also helps to reduce the number of interpretive options left open by the holism of the
mental. And contrary to Davidson’s claim, we can make perfectly good sense of a
distinction between generality and conjoined particularity in non-linguistic behaviors.

The *extensionality requirement* can be met by appeal to intrasubjective
triangulation. Brandom’s requirement that the being engage in manipulation of subject
terms is extravagant. And divergent causal chains leading from a worldly object to
multiple sensory modalities within a lone subject can play the role Davidson assigns to
multiple subjects in the determination of an object of thought.
The *reflection requirement* is not legitimate. Practical reasoning can be modeled in ways that do not render it reflective. We need not ever be aware of our beliefs as such in order for them to play the distinctive role of beliefs in our cognition. Nor do we need to consciously reflect on our mental representations of particulars in order to form general concepts.

The *control requirement* does not rehabilitate the reflection requirement, either. Deliberate action requires a being to pause to entertain various propositions to which it in fact takes certain attitudes. The Fregean complaint that a non-linguistic being would have no medium in which to entertain propositions is based on the mistaken notion that entertaining a proposition must be akin to observing it. Interpreters can seek the rudiments of control in subjects by attending to whether the subjects are capable of delaying action, and whether the subjects’ behavioral patterns are malleable with respect to similar stimuli. A non-reflective deliberation requirement of this kind captures the legitimate motivations behind the control requirement, and replaces the poorly motivated reflection requirement altogether.

The *objectivity requirement* is legitimate. I have argued against all reflection-based motivations for holding that only beings who grasp the distinction between subjective and objective can be in cognitive states that bear true or false contents. And I have also argued that an inkling of the distinction between subjective and objective could emerge from instances of intrasubjective triangulation in which one sense prompts the being to behave in a way that conflicts with the prompts from other senses, from the role of sentience-related phenomena like attention in the reasoning process, and from a sense
of bodily self that derives from proprioception. These arguments disarm most of the motivations for thinking that only linguistic beings could meet the objectivity requirement.

Arguments against the ability of non-linguistic subjects to meet the objectivity requirement cannot be redeemed by appeal to the institution requirement, either. The demand that a non-linguistic being’s cognitive processes must be held accountable to standards of truth and rationality can be met as long as the being has gone through a learning process in which its cognitive system has been modified for the purpose of bringing its activities more closely in line with the standards set by truth and rationality. The justification for claiming that the modifications made during the learning process are governed by the standards of truth and rationality derives from the fact that survival and flourishing can be treated as default ends for any creature, and learning aids a creature in meeting these ends because it brings the creature’s cognitive activities into closer approximation to the standards set by truth and rationality. Either the learning process counts as a practice, in which case the institution requirement is legitimate and non-linguistic learners can meet it, or it does not count as a practice, in which case my account of the manner in which the learning process renders a being accountable to the standards of truth and rationality shows that the institution requirement is illegitimate.

Finally, the requirement of conceptual standards is met by a more complex version of the appeal to the learning process. Beginning with a set of concepts the subject might plausibly possess, given its perceptual sensitivities and learning environment, the simplest standards for concept application that demarcate the categories
that are most relevant to the subject’s flourishing, and concerning which the subject’s discriminatory capacities have been honed, are those to which the subject is accountable. The account ties up the remaining loose ends concerning intensionality and objectivity by showing that appeal to the connections between the motivations for each of these requirements and the requirement of concept application does not provide any additional reason for thinking that non-linguistic beings cannot meet those requirements. And it ties up the remaining loose ends concerning institution and reflection by showing that appeal to connections between the motivations for each of those proposed requirements and the requirement of concept application does not offer any new reason to think that those requirements are legitimate.

Thus, these investigations have yielded an account of the genuine requirements for proper interpretation of a subject as a thinker. This account draws a principled distinction between genuine thinkers and beings whose actions are merely describable in interpretive language in a way that does not deny the possibility of non-linguistic thinkers. A subject is a genuine thinker just in case: (a) it has been through a learning process in which its cognitive system was modified to better track highly flourishing-relevant categories and to better conform to the standards of truth and rationality, (b) it has at least two sensory modalities, (c) it sometimes behaves in ways that are not immediate, invariant, and rigid responses to stimuli, but rather indicate the ability to pause to entertain various propositions, and (d) it displays an overall pattern of cognitive and behavioral activity in which a rational pattern can be found. In order to properly attribute thoughts to a non-linguistic subject that meets these conditions, an interpreter
should begin by paring down from her own set of concepts to those that demarcate
categories that are highly relevant to the subject’s flourishing, to which the subject is
discriminatorily sensitive (though not necessarily perfectly so), and for which this
sensitivity has been honed through exposure to members during the learning process.
From this point on, the process is guided in the same way as interpretation of a linguistic
subject: the interpreter strives to identify the simplest pattern of attributions that render
the subject’s activities readily explicable in rational terms on an activity-by-activity basis,
and highly rational overall.

In Chapter 5, I will provide further clarification of (i) the manner in which this
account fits within the interpretationist approach I have developed, (ii) the nature and
metaphysical determination of thought according to that approach, and (iii) the
implications of my account for the scientific, philosophical, and everyday approaches to
thought attribution.
Chapter Five: Interpretationism and Thought Without Language

Introduction

In the first four chapters, I have accomplished my goals in this dissertation. In Chapter 2, I argued for a non-scientific interpretationism about the thinking mind. In Chapter 3, I identified seven proposed requirements that a creature must meet in order to be a thinker, and considered the various arguments that have been and could be constructed against the possibility of non-linguistic thinkers on the basis of those proposed requirements. In Chapter 4, I demonstrated that none of those arguments is sound, and developed an account of the genuine requirements that a creature must meet in order to be a thinker. My account meets the demands I put forward in Chapter 1: it draws a principled line between thinkers and non-thinkers, and it does not restrict thought to linguistic beings.

This chapter will provide further clarification on some of the key matters that have been discussed in previous chapters. For one, I want to clarify the manner in which thoughts are metaphysically determined on my view, in order to better assess the extent to which my account preserves everyday and scientific attributive practice. I also want to make it clear that, although I have leaned on the work of Dennett and Davidson in my
development of my view, my view differs from each of theirs in some important ways. Additionally, I believe that some of the theoretical commitments I have undertaken in the course of developing my view of thought suggest directions for further theorizing in both metaphysics and value theory.

The Thinker/Non-thinker Distinction on my Approach and in Real Practice

On the account I offer, a subject counts as a thinker just in case:

(a) it has been through a learning process in which its cognitive system was modified to better track highly flourishing-relevant categories and to better conform to the standards of truth and rationality;
(b) it has at least two sensory modalities;
(c) it sometimes behaves in ways that are not immediate, invariant, and rigid responses to stimuli, but rather indicate the ability to pause to entertain various propositions;
(d) it displays an overall pattern of cognitive and behavioral activity in which a rational pattern can be found.

Conditions (a), (b), and (c) allow me to draw a principled distinction between thinkers and non-thinkers, because they embody constraints needed to justify the claim that a subject is accountable to rational standards, and not merely describable in rational terms. Such a distinction cannot be drawn on an account that appeals only to considerations like (d). Other philosophers have been willing to appeal to accounts that do not go much further than (d). Kripke allows us to attribute thoughts as long as the subject passes our usual, everyday tests for attribution. Burge allows us to attribute thoughts to non-linguistic animals as long as their behaviors fit the rule for concept application. Dennett,
even in his more nuanced, post-“Real Patterns” approach, allows us to seriously attribute thoughts as long as doing so allows us to make explanatory progress beyond what can be accomplished by taking a lower-level stance with respect to the subject. Room can be made on any of those approaches for non-linguistic thinkers, but at too great a cost, in my judgment. Davidson’s complaint that rationalizing explanations are far too easy to come by, and can be given for such movements as the ones a missile makes when it is fired toward a target, demonstrates a need to move well beyond (d), and to find a way to distinguish, among those items whose movements meet (d), the ones that are really accountable to rational standards and the ones that are not.

The primary achievement of my account is that conditions (a), (b), and (c) allow us to sort through the items whose movements meet (d) ---and given Davidson’s point about the missile, that class probably includes every item that moves. My account is certainly not the only one that does so. Davidson and Brandom both offer accounts on which a principled distinction between thinkers and non-thinkers can be drawn. But their accounts run contrary to current scientific and everyday attributive practice, in that the manner in which they draw the distinction between thinkers and non-thinkers rules out non-linguistic beings from counting as thinkers.

My view does not rule out non-linguistic thinkers, as can be shown by considering again Weiss’s character Arthur. Arthur clearly meets conditions (a) and (c). His experiences with metals have yielded an ability to craft them with fire and hammer, and he apparently finds some utility in that activity. He clearly does not respond in a pre-

programmed, inflexible way to the new metal that has a higher melting point than any he has encountered before. It would appear that Arthur meets (b) as well. His smith’s work would be difficult to carry out if he could not see the metal, feel the fire, or something comparable. We have every reason to believe that Arthur meets all of my conditions for counting as a thinker. And Arthur is stipulated to be non-linguistic.

So, my view is clearly compatible with the possible existence of non-linguistic thinkers. Though I do believe this result represents an advance over accounts that render non-linguistic thinkers impossible, it does not really constitute victory given my motivations. My primary complaint against views like Davidson’s is that their refusal to acknowledge non-linguistic thinkers embodies a divergence from everyday and scientific attributive practice, in which thoughts are attributed to non-linguistic beings. But the demonstration that my account allows for Weiss’s non-linguistic science fiction character to count as a thinker does not constitute a demonstration of convergence between my account and everyday and scientific attributive practices. It merely shows that there is one less barrier to such convergence on my account than on Davidson’s or Brandom’s.

The question I would really like to answer affirmatively is, “Does my account allow more or less the same non-linguistic beings to count as thinkers that everyday and scientific attributive practice count as thinkers?”

In response to this question, I think I succeed at least in ruling out many of the same items from counting as thinkers that are ruled out in those practices. For example, items such as thermostats and missiles fail to meet my minimal conditions for thinkerhood. Thermostats and missiles react immediately and invariantly to stimuli.
They do not undergo learning processes, nor do they flourish, as that involves healthy growth and development, and thermostats and missiles neither grow nor are apt for assessment as healthy or unhealthy. It is unclear that they could meet the rational standard of utility maximization, since that would require something to be useful to them, and it is unclear why one would think a thermostat could have ends to be met.\(^{154}\)

Very simple creatures whose behaviors are all automatic, invariant responses to stimuli will also fail to count as thinkers on my view. I suppose that the vertebrates, and even some of the “higher” invertebrates, could arguably meet the conditions for thought on my account. Routley’s description of the rufous fantail bird, for example, seems to fit my account of thought.\(^{155}\) It examines several potential nest sites before beginning to build in one that appears to be well-hidden from predators, showing an ability to pause and consider alternatives before acting. And when its nest is disturbed, it abandons the nest and seeks out a more favorable location before building a new one, showing an ability to modify its behaviors in ways that better accomplish its goal of building a nest that is safe from predators. And of course, the bird has multiple sensory modalities.

Hare’s chimps meet my conditions, as well.\(^{156}\) Along with the obvious multiple sensory modalities, the chimps display an ability to modify behavioral patterns in utility-maximizing ways in response to the rank and epistemic status of their peers. When placed in an enclosure baited with food, normally only the dominant chimp will go after

\(^{154}\) But suppose one had a very fancy thermostat, one that could take in information in multiple ways, correct its own tendencies to malfunction, pause before initiating changes in temperature, grow or in some sense count as prospering to various extents, with some states of affairs counting as more useful to it than others. In that case, one would have a thinking thermostat. The thing on my apartment wall does not even come close to meeting these conditions.

\(^{155}\) Routley pp. 399-400.

\(^{156}\) Hare et. al. (2000).
the food. But in cases in which the subordinate chimp has observed that the dominant
chimp cannot see the food and was not able to witness the baiting, the subordinate will
attempt to take the food unbeknownst to the dominant. This shows not only an ability to
modify behavior, but also an ability to attend to and respond to the cognitive states of
others, suggesting not only thought, but perhaps higher order thought.

Computers and any other form of artificial intelligence will be able to count as
thinkers on my view only if something fills the role of sensory organs, and something
fills the role played by the default ends of survival and flourishing in my account of non-
linguistic thinkers. Otherwise, they will not be able to engage in the intrasubjective form
of triangulation that I appeal to in order to meet the extensionality requirement, and their
activities will not be clearly rule-assessable rather than just rule-describable.157 I think all
of these facts accord well with everyday interpretive practice. Outside of the scientific
study of artificial intelligence, at least, it accords well with scientific attributive practice,
too. In each case, attributions are sometimes made to beings at the level of a fish or
octopus, but one would not want to go much further than that.

There is a further issue with the comparison of my account’s results to those of
contemporary scientific study of thought, however. The issue is that it is often unclear
that the “thoughts” that are attributed within scientific practice are really supposed to be
propositional attitudes, and that the process in which they are manipulated is really
supposed to be reasoning, rather than navigation via mental maps or some other form of
cognition that bears little structural/normative resemblance to language in the first place.

157 Similar considerations will apply to attributions of intentionality to groups or collections of individuals,
but in this case, it seems to me that many groups will meet the standards set by my account.
The worry here is not just that it will be difficult to tell how closely my account corresponds to scientific attributive practice. It is that science might not actually support my view that there can be thought without language, after all.

I do not mind this result, however. I have argued my case from the point of view of a non-scientific interpretationist. I would like to be able to claim that my account corresponds closely with scientific attributive practice, but I argued in Chapter 2 that it need not do so in order to claim to identify a real structure in cognitive processes. If it turns out that contemporary scientists are tracking different patterns in those processes, that is fine with me. Perhaps it even makes my project more interesting, in that I would then be arguing from a minority view within academia broadly construed, and not just from a minority view within the philosophical study of thought. I think that everyday practice still motivates a concern for drawing a principled distinction between thinkers and non-thinkers such that non-linguistic creatures can count as thinkers, regardless of whether scientific practice is in concurrence. Ultimately, I am willing to let my account stand on its own merits. I believe that my account provides support for our contemporary scientific attributive practice, in addition to the support it provides for everyday attributive practice, but if this turns out to be false, I am happy to argue as a revisionist.

It should be noted that one other factor makes comparisons between my account and our actual practices difficult. Our practices are carried out in the real world, where interpreters always have to work within the epistemic boundaries in which they find themselves. On my account, a creature’s history as a learner is important to proper interpretation of the creature, as is the creature’s complete set of dispositions to react to
new inputs with patterns of cognitive-behavioral responses. Clearly, this kind of
information will not usually be available to an interpreter, and when it is, that will
typically be the case only because the subject’s patterns of responses are too limited and
inflexible for the creature to meet my condition (c) for being a thinker. Though I have
offered guidelines for an interpreter to follow in order to apply my account, it should be
understood that the account I have offered is ultimately an account of the metaphysical
determiners of thinkerhood, and not an account of the practical workings of proper
interative process. Only under ideal epistemic conditions could these two be collapsed
into one, and real interpreters are never in ideal conditions. So, some divergences
between my account and real practice could be explained in terms of the epistemic
limitations of real interpreters. Real interpretive practice is going to have to license
attributions to be made in the absence of all the relevant information about metaphysical
determiners of thought. This is another reason, I think, not to follow the Kripkensteinian
proposal to treat everyday practice as if it provides all the metaphysical determination
that there can be of semantic facts. We can clearly see that there is room for a distinction
between what the practice must license if it is to operate at all and what the practice
would license under ideal circumstances. The latter, not the former, captures all that
there can be of semantic facts.

The Intentional Stance, Ideal Interpreters, and Indeterminacy
In another way, of course, I do think that everyday practice is authoritative. My non-scientific interpretationism commits me to thinking that everyday practice has the authority to establish that cognition has the structure of propositional attitudes/inference, independently of whether that structure is the one that best fits scientific criteria for “the” structure of cognition. Everyday practice finds this structure by attending to one kind of pattern found, in a rough and noisy manner, within cognitive events: the pattern defined by adherence to standards of rationality. But this irrevocable authority of everyday practice extends only over the basic structure of cognition as thought (i.e. inferences involving propositional attitudes). It does not extend to an authority over the specific contents of the thoughts involved in the generation of any specific action. These are fully determined only under ideal circumstances, and not by the justificatory standards for utterances of attributing sentences in everyday languages.

I do not think that my appeal to ideal circumstances threatens the considerations to which I have appealed in order to motivate interpretationism. I have argued that the metaphysical determination of thought must involve someone occupying the role of active interpreter. And I have not portrayed this merely as a matter of satisfying the need to find the structure of thought in the subject’s cognition, but also as a matter of deciding what pattern of attributions makes the best rational sense of the subject. And I have argued that the latter task requires someone to occupy an active role as a decision-maker.

I think that this aspect of my view is crucial to a strong defense of interpretationism against realism. There is a problem with arguing for interpretationism

---

158 I’m uncertain whether this commits me to the claim “All semantic truths are knowable”, though I suspect that the claim is true.
merely on the grounds that to judge cognition as thought is to judge it according to standards of rationality. A realist can respond by making a series of concessions to the epistemic necessity of taking up the intentional stance, and yet maintain realism. Yes, the realist may say, the theorist of thought needs to refer to rational standards in order to know what kind of patterns to look for in cognitive processes. But this is the way it works with every special science. One is never going to find facts about the workings of the food chain in a particular ecosystem, for example, by scanning through a complete particle-by-particle account of the physical system that embodies the ecosystem. One needs to know what type of patterns to look for in order to abstract from the physical data in the right way.\footnote{See Dennett (1991).} So, one would need to appeal to standards that specify what it is to be a predator, etc., in order to pick out the independently-determined facts about the workings of the food chain. Our need to take the intentional stance in order to appreciate the rational patterns in cognitive systems is no different, the realist might say. The appeal to the standards of the interpreter is just a preliminary step that must be taken in order to proceed to the task of finding the right interpretation-independent facts to ground realist type-identities or realization relations.

So, simply pointing out that one must take the intentional stance in order to carry out the task of thought attribution does not show that interpretationism is correct. A realist might even develop an atomistic theory of mental representation, and simply treat the interpretive process as an epistemic tool to use to check the correctness of one’s type identifications. If too much inexplicable disagreement occurs between one’s theory of

\footnote{See Dennett (1991).}
mental representation and interpretive practice, then this merely shows that one has not yet correctly specified the interpretation-independent identity relation. Furthermore, some disagreements might be explicable from the realist’s point of view simply by concluding that the interpretive practice is error-prone in various ways.

I am convinced that such a realist account is doomed to failure. Serious description of a subject in rational terms is partially a matter of choosing to hold the subject accountable to norms, and not merely a matter of describing it in terms of the standards to which it is objectively accountable, or the dispositions it actually has to behave, as realist accounts would have it. But I can imagine someone agreeing with my conditions on interpretation from the end of Chapter 4 and yet disagreeing with me that thought attributions essentially assess as well an describe. Because of this concern to show that attributions are not merely descriptive, I argued in Chapter 2 that the role of the interpreter must be an active role, and not one that could be filled simply by appeal to what an interpreter would say under such-and-such conditions. But this does create a difficulty for me, in that I want to appeal to idealized conditions in my account of the metaphysical determination of thought. And of course, an idealized interpreter cannot take an active role in anything, since it does not actually exist.

My solution is a fairly radical one. I claim that, in the absence of an omniscient interpreter (i.e. an interpreter that is ideal without needing to be idealized from non-omniscient interpreters), all thought is in fact indeterminate. My view of thought is a form of relativism. Semantic facts are fully determinate only relative to the contributions of an ideal interpreter. So, if there are no such beings, then the determiners of thought are
inadequate to complete their task. One essential component is missing: the decision of the active, epistemically ideal interpreter to treat the subject as being most rationally explicable according to one particular pattern of attributions.

When it comes to the matter of eliminating indeterminacies from thought, I do not insist that we can specify to uniqueness even with an epistemically ideal interpreter. For one thing, the utterances made or thoughts had by even an ideal interpreter in the course of giving an interpretation are themselves the type of things, linguistic or cognitive items, that need to be interpreted. Additionally, I do not think there is any fact of the matter whether, for example, a maximax or maximin strategy of choice under uncertainty is more reasonable, nor is there always going to be one clear way in which to resolve Quine-Duhem issues in cases where error must be attributed. So, there would be room for two ideal interpreters to disagree about which overall attribution pattern made more sense of a being. It is not just that there is a role for an ideal interpreter to play in the determination of thought; it is also that the role would have to be represented by a variable and not a constant value even for an all-evidence-in assessment of a subject’s overall pattern of activities. There is room for semantic truths to be multiplicitous over ideal interpreters’ decisions, not just relative to an ideal interpreter’s hypothetical decision.

To give a picturesque example, I think that even an ideal interpreter would be unclear as to “the” structure of Davidson’s Master Argument from “Rational Animals”. In this case, even the self-interpreter could not figure out exactly how best to bring together his intuitions into a compelling premise-by-premise presentation. I think this is
a dramatic example of indeterminacy at work; there simply is not a fact of the matter about how that argument was supposed to go. Let omniscience scan through every disposition Donald Davidson had, all the rules of term usage in the communities of English-users of which Davidson was a member, etc. ---there just is not enough there to specify one interpretation as “the” correct one. And though it is not always as dramatically apparent, I think this is the way things always are with thoughts and the sentences of natural languages. And though I have not argued for this claim, I think it is a natural development of the arguments I gave for interpretationism in Chapter 2, along with my unwillingness to ground a defense of interpretation on the epistemic necessity of taking up the intentional stance in order to find the structure of thought in cognition.

Incidentally, the maximax vs. maximin issue also explains my reluctance to think that Davidson’s appeal to Ramseyan decision theory is actually going to successfully pare down attribution patterns to uniqueness. I can appreciate Davidson’s desire to appeal to formal decision theory. It provides details about how an interpreter would go about finding rational patterns in behavior that go beyond those imposed by a principle of charity and the Tarskian Convention T. The decision-theoretical machinery should be regarded as filling in more details of the process of interpretation. But it should not change the outcome of the process. If the choice of one attribution pattern rendered the subject as playing a strictly dominant strategy, then that would have emerged in an idealized scenario where all behavioral evidence had been accounted for, whether or not one recognized that the interpreter’s process fit a decision theoretical model. On the other hand, if we have a case in which multiple interpretations remain equally viable even
in the idealized scenario, then this will be reflected in the decision-theoretic model by the availability of multiple strategies that offer the same overall expected utility. The decision-theoretical apparatus will not be able to identify any of these as most rational; appeal will need to be made to some choice strategy for resolving these deadlocks---maximin, for example. And I cannot see how one could claim to have a rational basis for choice of deadlock-breaking strategy. One might try to go descriptive here: which strategy does the subject actually seem to play? But in an idealized scenario, that would already be part of the data that led to the deadlock and not an additional bit of information to be appealed to in order to break the deadlock. A subject’s choice of deadlock-breaking strategy is part of an overall interpretation of the subject.

Neuroscience, the Metaphysical Determination of Thought, and the Motivations for Interpretationism

I think that my interpretationism differs from Davidson’s in another important way. In my version, the interpreter finds rational patterns in the overall cognitive and behavioral activities of the subject, and not merely by looking to the subject’s behaviors. I do not limit my account to the evidence that is available to a radical interpreter. This move is open to me in a way that it is not open to Davidson because I motivate interpretationism differently than he does. For Davidson, interpretationism is motivated in part by a desire to explain how each linguistic human begins life in a state of ignorance with respect to semantics, and ends up knowing how to use a language. He takes that to
indicate that all facts about meanings are accessible to the language learner, and the language learner has no tools to employ in this task beyond those of a radical interpreter.

But the inference from the nature of the process in which beings become linguistically competent to the conclusion that all semantically relevant factors are available in principle to a radical interpreter only goes through if we also assume that a being’s competence with a language entails that all factors relevant to the language’s semantics are accessible to the being. For most purposes, this is “true enough”, and there is no problem with claiming that the competent language user in effect knows a theory of truth for the language. But for the purpose of grounding a theoretical approach to the thinking mind, such considerations will not do. The theory of truth is known only in effect, and the semantically relevant features are often only as accessible as they need to be for us to get by. Learning to use a first language is largely a matter of learning a practical skill, and to describe the process as a matter of learning a theory of any kind can be little more than an approximation of what actually occurs.

In fact, I think that many semantically relevant factors will be beyond the competent language user’s rational ken. There are many reasons for this. There are Putnam’s considerations concerning the division of linguistic labor and Twin Earth cases. There is Burge’s discussion of the patient whose behavior displays an incomplete grasp of ARTHRITIS, and the identical behavior of the patient’s counterpart in a different community that displays mastery of some other concept. And there are events that impact the development of the brain and nervous system in ways that are relevant to
rational processes despite the fact that those events are not themselves amenable to characterization in terms of rationality.

For example, suppose I see a plant with flowers. I form a perceptual belief attributable as “There is a plant with flowers nearby.” Though I do not form perceptual beliefs regarding the leaf shape, flower shape, relative size of the flowers to leaves, etc., that information nonetheless reaches my brain via my perceptions. In some cases, we might want to say on the basis of my intake of information that I formed subconscious perceptual beliefs, but in other cases it will simply be implausible to hold that even this has happened. A complete novice when it comes to flowering plants might have so little inclination to care about leaf and flower size and shape that he could not even recall the information he took in about these aspects of the plant he saw. The information still made its way into his brain, though, and it might become accessible at a later time if, say, he suddenly found himself dating a nursery manager. He might learn more quickly that a particular combination of leaf and flower size and shape indicates that a plant is a begonia if he has been through repeated situations in which he has seen begonias, even if he has never formed even subconscious beliefs about the size and shape of the flowers and leaves of the plants he has seen. The sub-rational brain activity that occurred during those experiences might accumulate into an effect that is describable in rational terms.

Because of the intimate, entangled nature of the relationship between the rational thought process and the brain and phenomenal processes that help to give it its structure, I think it is foolish to restrict an account of the determination of thoughts and their contents to factors that are available to a radical interpreter. World-to-nervous system sensory
processes have aspects that will not figure in any noticeable way into rationalizing explanations of the concurrent activities of the subject, but will nonetheless play a causal role in shaping the underlying brain structures that realize each token thought.

Interpretation is a matter of making rational sense of a subject, and some facts that are not accessible to a radical interpreter can help to make better sense of a subject. If a certain type of structure is required in order for information to be stored as memories, and a being lacks that type of structure, we can conclude that any delays between stimulus and the being’s response are not indications that the being is pausing to entertain various propositions before acting---with no memory, the being cannot work offline on a problem. Study of the brain can help us to better understand which superficially similar behaviors are actually capable of constituting thought.

This manner of appealing to the nature of the brain does not yield realism about the thinking mind. On the basis of the considerations I discussed in Chapter 2, what I think we can never have is an atomistic theory of mental representation in which brain structures are type-identified with representations of specific contents. I think that any attempt to give such identifications would be question-begging, given that the method one must use to figure out the assignments is the holistic method of interpretation. A dogmatic realist could give such an account, but would have no way of providing any defense for it. So, anomalous monism is not threatened by any form of realist reduction simply by granting that certain brain structures can be type-identified with certain very broad mental functions. If it could be discovered that certain brain structures are required in order for a being to entertain a thought, my account could be supplemented with the
information that only beings with such structures could meet my deliberation
requirement. Such a discovery would not threaten my account; indeed, it should be
welcomed by a proponent of my account. Facts learned outside the practice of
interpretation should be applied to the practice’s conclusions, even if the ways in which
those facts are learned cannot themselves become part of the practice.

I think it is important to see that an account like mine can welcome input from the
scientific study of the nervous system. My view is a non-scientific interpretationism, not
an anti-scientific interpretationism. I think this fact must be stressed, given that it is not
the case on the Davidsonian radical interpretationist view. One who is accustomed to
dealing with Davidson’s version of interpretationism might suspect that interpretationism
is essentially limited to a behavioristic evidence base. It is true that terms like “thought”
and “belief” exist because we want a vocabulary that allows us to talk about what is
going on when we use language/interact with the world, despite our lack of easy access to
facts about the neurological structures and processes that are causally responsible for
such activities. But the focus on explaining our behaviors in ways that primarily appeal
to evidence other than that discovered through neuroscience does not entail the
behavioristic conclusion that what goes on outside the realm of easy access is irrelevant
to proper interpretation.

General Metaphysics and an Ethics of Flourishing

234
Although my primary focus has been on the metaphysics of thought and its relationship to standards of rationality, I believe that the considerations to which I have appealed suggest further directions for research in metaphysics and value theory more generally construed. In the realm of metaphysics, I have suggested (in Chapter 2) that a proliferation of divergent potential assignments of pattern-to-noise is not specific to cognition, but is a fact that holds with respect to any phenomenon. But I have also argued the mere presence of a pattern does not indicate the reality of the kind of item the pattern embodies. I have claimed that only assignments of pattern-to-noise that contribute to systematic accounts of the phenomenon in question, and best serve either scientific purposes or some vital non-scientific purpose should be considered to justify a claim that a type of thing exists. I think the other obvious option would be a wildly inclusive view of reality, on which any kind of pattern that can be found can be considered to justify a claim that a type of thing exists. I am not altogether unsympathetic to such a view. I did not want to rely on it, however, in my defense of interpretationism against eliminativism and realism.

In value theory, I think that the recognition that survival and flourishing can be treated as default aims for any creature has interesting consequences for standards other than those of rationality. The appeal to flourishing suggests that one might explore a (very) roughly Aristotelian approach to ethical standards, in which healthy existence is treated as providing the basis for all moral value. Construction of a theory of accountability to moral standards upon this axiological basis cannot be carried out as straightforwardly as my theory of accountability to standards of truth, rationality, and
concept application, because the relationship between what is of value to individuals and what is moral is not as straightforward as the relationship between what is of value to individuals and what is rational. So, I would not expect a theory of accountability to moral standards that allowed non-linguistic beings to count as moral agents to be forthcoming. But the appeal to flourishing could nonetheless provide an interesting constraint on ethical theorizing.

The Effects of Linguistic Practice on Accountability to Conceptual Standards

Because I have been mainly concerned with establishing the possibility of non-linguistic thinkers, I have not spent much time discussing the manner in which interpretation of linguistic beings is accomplished on my view. I think Burge, Putnam, and Davidson have already said much of what I would have wanted to say, anyway. The interpretation of linguistic beings is more complex than the interpretation of non-linguistic beings because one is driven to finer-grained interpretations. Furthermore, one must be aware not only of the linguistic subject and its natural environment, but also the standards of its linguistic communities. Typically, the appeal I make to the default ends of survival and flourishing to set conceptual standards in my account of a minimal requirement for thought will not be relevant to interpretation of linguistic beings. The instrument of language allows linguistic beings to think about items in ways that bear no straightforward relationship to the beings’ default interests. Because everyday practice treats language as an instrument of thought expression, the relatively stable and publicly
accessible standards for term use tend to take over the role of determining to what conceptual standards a being’s thoughts are to be held accountable.

This has an interesting result for one of the more puzzling subjects of interpretation ever to be imagined, Davidson’s Swampman.\textsuperscript{160} Swampman is, of course, a physical replica of Davidson, assembled in a swamp as the result of an extremely unlikely natural coincidence. Does Swampman have thoughts, on my account? One might think not, as he has not been through any kind of learning process ensuring that his cognitive activities have become governed by any set of conceptual standards. On the other hand, he has the full set of dispositions of a user of English, so the conceptual standards of the utterances he makes, construed as spoken English words, ought to govern his thoughts, as well. This complication would not exist for, say, Swampdog, a similarly-formed replica of a dog, who clearly will not be accountable to any standards initially on my view, and so will not be a thinker at first. It is tempting to say that Swampman, too, will be unaccountable to any standards upon his appearance---he has no history of interactions that would render him beholden to those English standards to which his activities would conform every bit as well as Davidson’s. He does, however, have a standing disposition to self-interpret as an English speaker, and dispositions to defer to experts in the use of various English terms. Ultimately, Swampman is bound by the standards of English, and is therefore a thinker, because he is readily interpretable as acknowledging those standards as binding.

\textsuperscript{160} Davidson (1987).
Conclusion

I will concede that the considerations I have offered here do not constitute a *disproof* of the Linguistic Priority Thesis. I have not proved that non-linguistic beings are capable of pausing to entertain various propositions before carrying out a behavioral process, and I would have to prove that in order to claim that language is not somehow essentially involved in such a process. Nonetheless, I have demonstrated that the Linguistic Priority Thesis lacks sufficient motivation to warrant our allegiance to it. Instead of a proof that non-linguistic beings can pause to entertain propositions, I have argued that there is no good reason to believe that the ability to entertain a proposition is exemplified by, or akin to, observing written characters or spoken words. And I have shown how the considerations that motivate the Linguistic Priority Thesis are compatible with the existence of non-linguistic thinkers, once they are properly formulated as the list of legitimate requirements for attribution of thought to a subject.

My account provides a workable alternative to those on which the Linguistic Priority Thesis holds. It incorporates the considerations that motivate the Linguistic Priority Thesis without yielding a commitment to LPT itself. Because my account draws a principled distinction between thinkers and non-thinkers that allows some non-linguistic beings to be counted among the thinkers, there is hope of synchronicity between my account and everyday and scientific attributive practices. The approaches of Linguistic Priority Theorists do not offer such hope. Given that my view holds this clear advantage over views that accept Linguistic Priority, and views that accept Linguistic
Priority hold no advantage over my view, we are warranted in rejecting the Linguistic Priority Thesis.
References


Ben-Yami, Hanoch. “Against Characterizing Mental States as Propositional Attitudes.” *Philosophical Quarterly* 47 (186): 84-89, Ja 97.


Davidson, Donald. “Comments on Karlovy Vary Papers”. In Interpreting Davidson, Petr Kotatko, Peter Pagin, and Gabriel Segal, eds. CSLI Publications, 2001.

Davidson, Donald. “Externalisms”. In Interpreting Davidson, Petr Kotatko, Peter Pagin, and Gabriel Segal, eds. CSLI Publications, 2001.


Frege, Gottlob. “Function and Concept”. Lecture to the Jena Society for Medicine and Natural Science, Jan 9, 1891.


Hintikka, Jaakko. “Sosa on Propositional Attitudes *De Dicto* and *De Re*.” *Journal of Philosophy* 68: 489-497, Ag 1971.


257


Appendix: the ten arguments and the points at which they invoke the seven purported requirements

1 Davidsonian Holism Argument:

1 If Ramseyan methods cannot be applied, then the holism of the mental leaves open an unacceptably large number of competing interpretations. (P → Q)
2 If we cannot judge subjects’ selections at the grain of intensions, then we cannot apply Ramseyan methods. (R → P)
3 If a subject is non-linguistic, then we cannot judge the subject’s selections at the grain of intensions. (S → R)
4 So, if a subject is non-linguistic, then the holism of the mental leaves open an unacceptably large number of competing interpretations. (S → Q)


The Intensionality Requirement is motivated at line 2 and connected to language use at line 3.

2 Davidsonian Network Argument:

1 The specific conceptual composition of a thought is determined by its place in a dense, extensive network of beliefs. (P ↔ Q)
2 Only linguistic subjects behave in ways rich enough to embody a dense, extensive network of beliefs. (Q → R)
3 So, only linguistic subjects can have thoughts with determinate conceptual contents. (P → R)

Assume P for conditional introduction. P, P ↔ Q ├ Q. Q, Q → R ├ R.

The Intensionality Requirement is invoked at line 1.

3 Davidsonian Error Argument Schema:
1 Belief is only made intelligible in contexts involving error. (P → Q)
2 The contrast between error and correct conduct exists, or is made intelligible, only in a norm-governed situation (in the case of belief, the relevant norm is set by the standard of objective truth). (Q → R)
3 Objective truth exists, or is made intelligible, only in contrast to the held true. (R → S)
4 The notion of being held true applies only to sentences, or is made intelligible only in the context of its application to an interpretee’s sentences. (S → T)
5 Only beings to whom belief is intelligible can have a belief. (U → P)
6 A being must have beliefs in order to have any thoughts. (V → U)
7 So, only linguistic interpreters can have thoughts. (V → T)


The Institution Requirement is invoked at line 2, pre-parenthesis. The Objectivity Requirement is invoked at line 2 within the parentheses. The Reflection Requirement is invoked at line 5.

4 Davidsonian Concept Cluster Argument:

1 The concepts BELIEF, TRUTH, and ERROR are acquired only as a group. (P → Q)
2 That group of concepts is acquired only in the process of linguistic interpretation. (Q → R)
3 Grasp of the concepts BELIEF, TRUTH, and ERROR is required in order for a being to be in states that function as beliefs. (S → P)
4 A being must have beliefs in order to have any thoughts. (T → S)
5 So, only linguistic interpreters can have thoughts. (T → R)

Assume T for conditional introduction. T, T → S ⊢ S. S, S → P ⊢ P. P, P → Q ⊢ Q. Q, Q → R ⊢ R.

The Reflection Requirement is invoked at line 3.

5 Davidsonian Generality Argument:

1 The specific conceptual composition of a thought is determined by its place in a network of beliefs. (P ↔ Q)
2 The network must contain general beliefs if it is to identify the specific conceptual composition of a thought. (Q → R)
3 Only when a subject uses categorical terms do we have any way to make sense of the distinction between general beliefs and conjunctions of beliefs about particulars. (R → S)
4 So, only linguistic subjects (i.e., subjects who use categorical terms) can have thoughts with determinate conceptual contents. (P → S)


The Intensionality Requirement is invoked at line 1.

6a1 Davidsonian Master Argument Subargument 1, version 1:

1 To have a belief that one could readily discover to be false entails that one is potentially subject to surprise. (P → Q)
2 One cannot be a thinker without having some beliefs that one could readily discover to be false. (R → P)
3 Surprise requires an awareness of a contrast between something that was believed and something one has come to believe. (Q → S)
4 Awareness of a contrast between something that was believed and something one has come to believe requires grasp of the concept of belief. (S → T)
5 So, to be a thinker, one must have the concept of belief. (R → T)


The Reflection Requirement is invoked at line 4.

6a2 Davidsonian Master Argument Subargument 1, version 2:

1 Reflective surprise requires one to have beliefs about the correctness of one’s beliefs. (P → Q)
2 To have beliefs about the correctness of one’s beliefs requires that one grasps the concept of belief. (Q → R)
3 To have a general stock of beliefs, one must be subject to reflective surprise. (S → P)
4 To be a thinker, one must have a general stock of beliefs. (T → S)
5 To be a thinker, one must have the concept of belief. (T → R)

Assume T for conditional introduction. T, T → S ⊢ S. S, S → P ⊢ P. P, P → Q ⊢ Q. Q, Q → R ⊢ R.

The Reflection Requirement is invoked at line 3.
6a3 Davidsonian Master Argument Subargument 1, version 3:

1. Thinkers must be able to take deliberate control of their actions.  \((P \rightarrow Q)\)
2. In order to take deliberate control of their actions, beings must consider how their beliefs advise them to act in the pursuit of their goals.  \((Q \rightarrow R)\)
3. In order to consider how their beliefs advise them, beings must be able to focus their attention on their beliefs.  \((R \rightarrow S)\)
4. In order to focus attention on one’s beliefs, one must be able to conceive of one’s beliefs as such.  \((S \rightarrow T)\)
5. Thus, thinkers must be able to conceive of their beliefs as such.  \((P \rightarrow T)\)

Assume \(P\) for conditional introduction.  \(P, P \rightarrow Q \vdash Q\).  \(Q, Q \rightarrow R \vdash R\).  \(R, R \rightarrow S \vdash S\).  \(S, S \rightarrow T \vdash T\).

The Control Requirement is invoked at line 1.  The Reflection Requirement is invoked at line 4.

6b1 Davidsonian Master Argument Subargument 2, version 1:

1. Only through grasp of the concept of intersubjective truth could one come to grasp the subjective-objective contrast.  \((P \rightarrow Q)\)
2. Grasp of the concept of intersubjective truth depends on participation in linguistic communication.  \((Q \rightarrow R)\)
3. A being can make sense of the dual concepts of subjective belief and objective truth only in their roles as embodiments of the contrast between subjective and objective.  \((S \rightarrow P)\)
4. So, only those who have communicated linguistically can grasp the concept of belief.  (from 3, 4)  \((S \rightarrow R)\)

Assume \(S\) for conditional introduction.  \(S, S \rightarrow P \vdash P\).  \(P, P \rightarrow Q \vdash Q\).  \(Q, Q \rightarrow R \vdash R\).

6b2 Davidsonian Master Argument Subargument 2, version 2:

1. Only a being that recognizes triangulation to be an intersubjective practice governed by a shared standard will be able to understand divergences from expectations as instances of error.  \((P \rightarrow Q)\)
2. A being must understand divergences from expectations as instances of error if it is to make sense of the notion of belief.  \((R \rightarrow P)\)
3. Only linguistic communication causes a being to understand that it shares a standard—intersubjective truth—with others.  \((Q \rightarrow S)\)
4 So, only linguistic communication suffices to bring a being to understanding of the notion of belief. \((R \rightarrow S)\)

Assume R for conditional introduction. \(R, R \rightarrow P \vdash P.\) \(P, P \rightarrow Q \vdash Q.\) \(Q, Q \rightarrow S \vdash S.\)

7 Basic Institutional Argument:

1 Thinkers’ behavioral and cognitive processes must be accountable to standards of truth, rationality, and concept application, not merely describable in terms of such standards. \((P \rightarrow Q)\)
2 Only beings who have been initiated into a practice are accountable to that practice’s standards (in this case, standards of truth, rationality, and concept application). \((Q \rightarrow R)\)
3 Only linguistic beings have been initiated into a practice governed by standards of truth, rationality, and concept application. \((R \rightarrow S)\)
4 So, only linguistic beings are thinkers. \((P \rightarrow S)\)

Assume P for conditional introduction. \(P, P \rightarrow Q \vdash Q.\) \(Q, Q \rightarrow R \vdash R.\) \(R, R \rightarrow S \vdash S.\)

The Objectivity Requirement is invoked at line 1. The Institution Requirement is invoked at line 2.

8 Kripkensteinian Argument:

1 Thinkers’ behavioral and cognitive processes must be constrained by standards of concept application. \((P \rightarrow Q)\)
2 Considered in isolation from linguistic practices, the activities of a subject fail to determine whether the subject is accountable to a standard of concept application that it meets, or to a standard it fails to meet. \((R \rightarrow S)\)
3 If it is indeterminate whether a subject’s activities are in accord with the standard of concept application to which it is accountable or in conflict with the standard to which it is accountable, then the subject is not constrained by standards of concept application. \((S \rightarrow \neg Q)\)
4 If a being was not initiated into linguistic practice, then it must be considered in isolation from linguistic practices. \((T \rightarrow R)\)
5 If a being is non-linguistic, then it was not initiated into linguistic practice. \((U \rightarrow T)\)
6 So, if a being is non-linguistic, then it is not a thinker. \((U \rightarrow \neg P)\)

Assume U for conditional introduction. \(U, U \rightarrow T \vdash T.\) \(T, T \rightarrow R \vdash R.\) \(R, R \rightarrow S \vdash S.\) \(S, S \rightarrow \neg Q \vdash \neg Q.\) \(\neg Q, P \rightarrow Q \vdash \neg P.\)

269
9 Brandomian Deliberation Argument:

1 Deliberation involves considering what commitments would be attributed by linguistic scorekeepers under various conditions. \((P \rightarrow Q)\)
2 Only beings who have become proficient at the skill of linguistic scorekeeping can consider what commitments would be attributed by linguistic scorekeepers under various conditions. \((Q \rightarrow R)\)
3 Only linguistic beings are proficient at the skill of linguistic scorekeeping. \((R \rightarrow S)\)
4 Deliberation is required for authority over one’s own behaviors and cognitive processes. \((T \rightarrow P)\)
5 To be accountable to rational standards, beings must have cognitive authority over their behaviors and cognitive processes. \((U \rightarrow T)\)
6 Thinkers as such are accountable to rational standards. \((V \rightarrow U)\)
7 So, only linguistic beings can be thinkers. \((V \rightarrow S)\)

Assume \(V\) for conditional introduction. \(V, V \rightarrow U \vdash U.\) \(U, U \rightarrow T \vdash T.\) \(T, T \rightarrow P \vdash P.\) \(P, P \rightarrow Q \vdash Q.\) \(Q, Q \rightarrow R \vdash R.\) \(R, R \rightarrow S \vdash S.\)

The Institution Requirement is invoked at line 1. The Reflection Requirement is invoked at line 4.

10 Brandomian Syntactic Priority Argument:

1 Thoughts must bear conceptual contents that represent a world of objects as having certain properties (i.e., to be a thinker, one must confront a world of propertied objects). \((P \rightarrow Q)\)
2 Only a being who participates in a linguistic practice with subject-predicate structure confronts a world of propertied objects. \((Q \rightarrow R)\)
3 So, only a being who participates in a linguistic practice with subject-predicate structure is capable of thought. \((P \rightarrow R)\)

Assume \(P\) for conditional introduction. \(P, P \rightarrow Q \vdash Q.\) \(Q, Q \rightarrow R \vdash R.\)

The Extensionality and Intensionality Requirements are invoked at line 1. The Institution Requirement is invoked at line 2.