The Inessential Indexical
CONTEXT AND CONTENT

Series editor: François Recanati, Institut Nicod

Other titles in the series:

The Mirror of the World
Subjects, Consciousness, and Self-Consciousness
Christopher Peacocke

Assessment Sensitivity
Relative Truth and its Applications
John MacFarlane
The Inessential Indexical

On the Philosophical Insignificance of Perspective and the First Person

Herman Cappelen
and Josh Dever
HC: For Nora
JD: For Brigid, Maeve, and Sophia
Acknowledgments

Philosophizing is best done in the company of others. We have been extraordinarily lucky in having a very large number of generous and insightful philosophical companions to help us think about the issues in this book. We are particularly grateful to three people. Andy Egan is as close to a co-author of this book as one can be without having written any of it and while disagreeing with most of it. He gave us detailed comments on every page of the book, many of which triggered radical changes along the way. Conversations with (and the work of) Dilip Ninan motivated our decision to write this book and helped us understand these extraordinarily difficult issues. Seth Yalcin gave constructive detailed comments on the entire manuscript and crucial critical feedback on and off over the last three years.

We are grateful for conversations with and feedback from: Pranav Anand, Derek Ball, Berit Brogaard, Ray Buchanan, David Chalmers, Cian Dorr, Olav Gjelsvik, Michael Glanzberg, Jeremy Goodman, Patrick Greenough, John Hawthorne, Torfinn Huvenes, Hans Kamp, Peter Ludlow, John MacFarlane, Ofra Magidor, Ruth Millikan, Daniel Morgan, Michael Murez, Anders Nes, John Perry, Simon Prosser, François Recanati, Susanna Schellenberg, Anders Schoubye, Jason Schukraft, John Searle, Susanna Siegel, David Sosa, Robert Stalnaker, Jason Stanley, Rachel Sterken, Andreas Stokke, Margot Strohminger, Sergio Tenenbaum, Mike Titelbaum, Michael Tye, Brian Weatherson, Timothy Williamson, and Juhani Yli-Vakkuri. Ephraim Glick provided insightful comments on the entire manuscript that helped us avoid mistakes and obscurities on several important points.

An early version of this material was presented at a conference organized by CSMN at the University of Oslo in June 2010. It was also presented as a series of lectures at Universidade Federal do Rio de Janeiro in March 2011, and then as a series of lectures at the Institut Jean Nicod in Paris in April 2011. Parts of this material were presented as talks in many locations around the world including Oxford, the University of Texas at Austin, the Arché Centre at the
University of St Andrews, and UC Berkeley. We are grateful to audiences for feedback on all these occasions.

Our research assistants, Margot Strohminger and Jason Schukraft, corrected an almost infinite number of typographical, editorial, and substantive mistakes in early versions of the manuscript.

Our editor, Peter Momtchiloff, generously provided encouragement, support, and advice during the three years we worked on this book.

Finally, we should highlight that most of our work on this book was made possible by two fantastic philosophical research centres, CSMN at the University of Oslo and Arché at the University of St Andrews.
## Contents

1. Introductory Overview: The Role of Indexicality, Perspective, and the De Se in Philosophy .............................................. 1

2. Preliminaries: Language-Mind, Super Indexicals, and Opacity ............................................................. 21

3. Indexicality, the De Se, and Agency ........................................... 30

4. Indexicality, Opacity, and Fregeanism ....................................... 58

5. Lewis on the De Se, Self-Ascription, and Centered Worlds ............... 86

6. Functionalism to the Rescue? .................................................. 109

7. Indexicality and Immunity to Error ............................................ 127

8. A Brief Note on Perceptual Content and the De Se ......................... 140

9. The De Se and the Semantics of PRO Constructions ................... 151

10. The View From Everywhere .................................................. 173

Bibliography ............................................................................. 183

Index ......................................................................................... 191
Detailed Contents

1. Introductory Overview: The Role of Indexicality, Perspective, and the De Se in Philosophy 1
   1.1. Perspective and Indexicality: Deep or Shallow? 1
   1.2. Our Target: The Almost Universally Accepted View that Indexicality Is Philosophically Deep 2
   1.3. On the Immense Impact of Essential Indexicality: Illustrative Quotes 6
   1.4. A More Systematic Presentation of Alleged Repercussions of Essential Indexicality Throughout Philosophy 9
   1.5. More Precise Articulation of Our Target: Essential Indexicality Theses and the Alleged Explanatory Roles of Indexicality 14
   1.6. Basic Kaplanianism and the Goals of This Book 15
   1.7. Some of the Many Issues We Will Not Address 17
   1.8. How to Read This Book 18

2. Preliminaries: Language-Mind, Super Indexicals, and Opacity 21
   2.1. Essential Indexicality: Language or Mind? 21
   2.2. Are All Indexicals Equal or Are There Super Indexicals? 26
   2.3. Opacity and Essential Indexicality 28

3. Indexicality, the De Se, and Agency 30
   3.1. Perry on the Opacity of Explanation Contexts 31
   3.2. Causation, Explanation, Motivation, and Rationalization 34
   3.3. Impersonal vs. Personal Action Explanations/Rationalizations 35
   3.4. Preliminaries: Some Weaker Theses and Why We Focus on IIC 37
   3.5. Motivation 1: Judgments about Cases 41
   3.6. Motivation 2: Bodily Movements Require Indexical Thought 42
   3.7. Motivation 3: Generalizations about Actions 45
   3.8. Motivation 4: Anti-Generalizations about Actions 47
   3.9. The Selection Problem and a Positive Proposal: The Action Inventory Model 49
   3.10. Clarifications, Objections, and Replies 52
   3.11. Brief Summary of This Chapter 57
4. Indexicality, Opacity, and Fregeanism
4.1. Puzzle Cases Involving Indexicals With Frege Counterparts 58
4.2. Some Irrelevant Cases 61
4.3. Two Confusing Cases: Prior’s “Thank Goodness That’s Over” and Lewis’s Two Gods 64
4.4. Preliminaries 1: Scope Restriction and Why Non-Fregeans Can Skip the Rest of This Chapter 66
4.5. Preliminaries 2: Robust Fregean Senses vs. Descriptive Fregean Senses 69
4.6. Claim 1: Fregeans Have a Problem With Indexicality, but It Has Nothing to Do With Opacity 71
4.7. Claim 2: Perry’s Famous Anti-Fregean Arguments Have Nothing Specifically to Do With Indexicality 72
4.8. Taking Stock 73
4.9. Objection 1: You Haven’t Refuted Fregeanism 77
4.10. Objection 2: You Haven’t Solved the Problem of Opacity 78
4.11. Objection 3: To Refute Essential Indexicality You Need a Theory of Content and to Have a Theory of Content You Need a Theory of Opacity 81
4.12. Brief Summary of This Chapter 82

5. Lewis on the De Se, Self-Ascription, and Centered Worlds
5.1. Lewis on Propositions and Properties 86
5.2. Response 1 to the Two Gods: Don’t Be Seduced by Metaphors 87
5.3. Response 2 to the Two Gods: Disentangling Propositions vs. Properties from De Dicto vs. De Re 91
5.4. Response 3 to the Two Gods: It’s Still All Frege Puzzles 95
5.5. Response 4 to the Two Gods: Worlds Are Not All There Is 99
5.6. The Role of Self-Ascription 103
5.7. Brief Summary of This Chapter 107

6. Functionalism to the Rescue?
6.1. Attack vs. Defense 109
6.2. Problems for Attack 110
6.3. Two Problems for Fancy Functionalism 111
6.4. Back to Defense 116
6.5. Brief Summary of This Chapter 120

7. Indexicality and Immunity to Error
7.1. Thesis 1: There is No Immunity to Error Through Misidentification 127
7.2. Thesis 2: Immunity Does Not Track Indexicality 133
7.3. Thesis 3: Indexicality Does Not Explain Immunity 135
7.4. Brief Summary of This Chapter 139

8. A Brief Note on Perceptual Content and the De Se 140
8.1. No New Arguments for De Se Contents 141
8.2. Addendum: Brogaard and Chalmers on Perceptual Content 146
8.3. Brief Summary of This Chapter 149

9. The De Se and the Semantics of PRO Constructions 151
9.1. Some Background: With Chierchia the De Se Moves From Philosophy to Linguistics 152
9.2. Some Linguistic Phenomena that Allegedly Require Distinctly De Se Interpretations 154
9.3. Our Central Concern: What’s the Philosophical Significance? 157
9.4. Brief Digression: Higginbotham on Self-Reflexivity 160
9.5. Challenging the Data 161
9.6. Warm-Up: Familiar Frege Puzzles 162
9.7. “Expects To” Without the De Se 163
9.8. “Admits To” Without the De Se 164
9.9. “Remembers” + Gerund Without the De Se 164
9.10. “Only”-Binding Without the De Se 165
9.11. Dream Reports Without the De Se 165
9.13. Brief Summary of This Chapter 171

10. The View From Everywhere 173
Chapter 1

Introductory Overview: The Role of Indexicality, Perspective, and the De Se in Philosophy

1.1. Perspective and Indexicality: Deep or Shallow?

When we represent the world in language, in thought, or in perception, we often represent it from a perspective. We say and think *that the meeting is happening now, that it is hot here, that I am in danger and not you; that the tree looks larger from my perspective than from yours*. The frequent perspectivality of representation can appear to be either a philosophically deep or a philosophically shallow phenomenon. To those who see it as deep, representational perspectivality results from a fundamental divide between representing the world objectively, as it really is, and representing it subjectively, as it seems to us. Perspectival representations are *representationally essential*: perspectivality makes available ways of representing the world that are not available with non-perspectival devices. And perspectival representations are *metarepresentationally essential*: many of the philosophical roles that we want representations to play can be played only by perspectival representations. The goal of inquiry is to determine which possible world is *actual*, which possible world *we* are in. Even stronger, for beliefs and desires to motivate us to action, they must not merely represent the world, but represent us as occupying a certain *position* in the world. Perception essentially presents the world as received from a *perspective*. 
But a competing view sees representational perspectivality as philosophically shallow. On this view, perspectival contents are just an instance of relational contents. Some representations depict how, e.g. New York is related to Chicago; others depict how the representer is related to Chicago. There is no deep difference in representational kind between the two. And there is no philosophically distinctive role to be played by perspectivality in the explanation of action, inquiry, or perception. There may be shallow contingencies here. Because our sense organs are close to us, we may happen to know more about how we are related to the world than about how other objects are related to the world, and because we are typically somewhat self-interested, we may care more about, and be more effectively motivated by, facts about how we relate to the world than by facts about how other objects relate to the world. But such shallow contingencies result from facts about our biology and psychology that don’t lead to a philosophically distinctive role for perspectival representation.

1.2. Our Target: The Almost Universally Accepted View that Indexicality Is Philosophically Deep

This book is an extended exploration and defense of the view that perspectivality is philosophically shallow. In doing so we are opposing one of the most entrenched and dominant trends in contemporary philosophy and maybe also in the history of philosophy: that perspectival (and the perspective of the first person in particular) is philosophically deep and that a proper understanding of it is important not just in the philosophies of language and mind, but throughout philosophy. Even though this view has important footholds throughout the history of western philosophy, the manifestation of it that will be most familiar to contemporary audiences is embedded in terminology and ideology introduced by John Perry and David Lewis. They introduced the labels essential indexicality (Perry 1979), irreducibly de se attitudes (Lewis 1979), and self-locating attitudes (Perry 1977). This work approaches these issues through a critical examination of the Perry–Lewis tradition.
Here is a provocative way to think of the goal of this book (more precise articulations are provided later): we argue that the terms just mentioned denote nothing—there is no such thing as essential indexicality, irreducibly *de se* attitudes, or self-locating attitudes. Our goal is not to show that we need to rethink these phenomena—that they should be explained in ways different from how, e.g. Lewis and Perry explained them. Our goal is to show that the entire topic is an illusion—there’s nothing there. The brilliant papers by Lewis and Perry lead people to believe in, and theorize about, an illusion. It isn’t that their theories fail to adequately explain the phenomenon they identified. The problem is deeper: there’s nothing to explain or account for. The way we see it, Perry, Lewis, and their followers have observed fragments of various different phenomena. They have then made the mistake of thinking that there’s something—call it essential indexicality—that ties all of these fragments together, when in fact there’s no unified phenomenon at all. It’s just a bunch of different phenomena with differing explanations.

We’ll start with a brief outline of this familiar terminological terrain. In his seminal 1979 paper, “The Problem of the Essential Indexical,” Perry introduced the terms “essential indexicality” and “locating beliefs.” He says, for example, “I shall use the term ‘locating beliefs’ to refer to one’s beliefs about where one is, when it is, and who one is. Such beliefs seem essentially indexical” (p. 5). What is it for a belief to be essentially indexical? Well, that, in a way, is the topic of this book. Here, for what it is worth, is the initial characterization we get from Perry: “When we replace it (the word ‘I’ in my expression of what I came to believe) with other designations of me, we no longer have an explanation of my behavior and so, it seems, no longer an attribution of the same belief. It seems to be an essential indexical” (p. 3). Note right away that it has something to do with action explanation and with substitutivity (or non-replaceability).1

David Lewis’s paper “Attitudes *De Dicto* and *De Se*” was published the same year as Perry’s paper “The Essential Indexical,” but Lewis (1979, p. 521) acknowledges a debt to earlier work by Perry (in particular Perry

---

1 Chs. 3 and 4 will explore these themes in more detail.
1977) and points out that Perry was indebted to Hector-Neri Castañeda. Lewis’s initial articulation of the view he thinks he shares with Perry and Castañeda is this: “Castañeda argues that the ‘he’ (or ‘he himself’) that appears for instance in “The editor of Soul knows that he is a millionaire’ is ineliminable. As I would put it, this typical attribution of knowledge de se is not equivalent to any attribution of knowledge de dicto” (pp. 521–522). The terminology of “attitudes de se” was introduced by Lewis in his paper and has now become common coin. We say more about how “the de se” can be understood (or more accurately, cannot be understood) later in this book, but for now it is worth noting that the term is introduced to mean more or less the same as what Perry meant by “essentially indexical attitude” or “self-locating attitude”—and this little cluster of terminology has shaped much of the contemporary discussions.

One way to get a pre-theoretic sense of what Lewis and Perry are after is through Perry’s little story of Lingens:

An amnesiac, Rudolf Lingens, is lost in the Stanford library. He reads a number of things in the library, including a biography of himself, and a detailed account of the library in which he is lost....He still won’t know who he is, and where he is, no matter how much knowledge he piles up, until that moment when he is ready to say, “This place is aisle five, floor six, of Main Library, Stanford. I am Rudolf Lingens.” (1977, p. 492)

Lewis says that what Lingens can’t get from books is irreducible de se beliefs. This is alleged to be the kind of insight into the world that is irreducibly perspectival.

The widespread endorsement of the Perry-Lewis view in contemporary philosophy is accompanied by an astonishing lack of critical literature. More or less the only anti-Perry-Lewis paper that has attracted some attention is Ruth Millikan’s 1990 paper, “The Myth of Essential Indexicality.” That paper is now twenty-three years old. The Perry-Lewis line has thus achieved an almost unique status in contemporary philosophy: It has become common ground among contemporary philosophers that (a) and (b) are true:  

---

2 Both Perry and Lewis acknowledge a debt to Castañeda’s work. Neither of them specifies in more detail what part of Castañeda they endorse or were inspired by.

3 For elaboration on this claim see Sect. 1.3 where we document this widespread endorsement in more detail.
a) Perry and Lewis have succeeded in articulating a clear thesis (or theses) and the terminology they introduced into the philosophical vernacular (“essential indexicality,” “irreducibly de se attitudes,” etc.) is well defined.

b) Perry and Lewis developed (at least) powerful and interesting arguments for the thesis of Essential Indexicality.

You can think of this book as having a Modest Goal and an Ambitious Goal:

**Modest Goal:** Throw significant doubt on both (a) and (b). We try to show that none of the published arguments in favor of Essential Indexicality are strong or convincing. We also show that the terminology introduced by Perry and Lewis is at best obscure and borderline unintelligible.  

**Ambitious Goal:** Prove that perspective and indexicality are philosophically shallow: they play no important explanatory roles in philosophy.

The ambitious goal is too ambitious to be fully achieved in a single work—it concerns foundational issues across large swaths of philosophy. No one or two philosophers can claim full expertise and competence in all those areas. And so we are relatively modest: although we think what we say here is an important step toward the ambitious goal, we think the answer is no. Our goal is primarily to look at what we have found explicitly in Perry, Lewis, and their followers and ask: Do those published claims add up to a reason to think indexicality is doing important explanatory work in philosophy? We will try to convince you that the answer is a resounding no.

---

4 We should note right away a methodological problem that faces projects like ours. When trying to show that those who use a certain term, e.g. “Blah,” are using a poorly defined term, it is often difficult to avoid using the dubious terminology. When we say what our opponents say, describe their arguments, and even in denying their views, it is tempting to use “Blah.” The cleanest option is to always use quotation marks and avoid using “Blah,” but that makes for very clumsy writing and so for ease of exposition we will throughout this work make use of our opponents’ terminology, but trust our readers to introduce quotation marks (or scare quotes or other distancing devices) where required.
1.3. On the Immense Impact of Essential Indexicality: Illustrative Quotes

It’s widely assumed that Essential Indexicality (and irreducibly *de se* attitudes) have important repercussions throughout philosophy. It is also widely assumed that it is widely assumed. These assumptions have jointly generated a philosophical environment in which it is acceptable to take as common ground that something properly labeled “Essential Indexicality” is true. In large swaths of contemporary philosophy this is the kind of assumption one can rely on without argumentation and without making clear what one means by it. It is as if the content of the thesis and the arguments for it are common knowledge. This is most easily seen in the now common way in which philosophers will make some rather vague and grandiose claim on behalf of indexicality (or first-person thought, or “the *de se*” or egocentric thought or what have you), and then casually add a footnote, “See Perry 1979 or Lewis 1979.” This section documents the immense impact of the Perry-Lewis view throughout contemporary philosophy.

Here is Shoemaker on the importance of “the first-person perspective” in philosophy of mind:

Some would say that the philosophy of mind without the first-person perspective, or the first-person point of view, is like *Hamlet* without the Prince of Denmark. Others would say that it is like *Hamlet* without the King of Denmark, or like *Othello* without Iago. I say both. (1994, p. 7)

Here is Bas van Fraassen on the role of indexicality in an account of scientific inquiry:

[A] view of science would hardly be empiricist if it ignored the *uses* of science, as a resource for praxis. How are theories and models drawn on to communicate information about what things are like, to guide our expectations in practical affairs, to design instruments and technological devices, to find our way around in the world? Here especially enters the indispensability of indexical judgment, even in contexts where the term “perspective” may not be apt. In the end, as we shall see, that entry of the essential indexical will prove to be the crucial clue to how science can depict at all. (2008, p. 88)
Here is Ruth Millikan describing what she took to be a widely endorsed view about agents’ ability to “project knowledge of the world into relevant action”:

A strong contemporary current runs to the effect that the ability of an agent to project knowledge of the world into relevant action in the world depends upon the ability to think indexical thoughts. For example, if I wish to get to Boston, it may be helpful to know that the 8:25 train goes there. But I cannot put this knowledge to use unless I also come to know, at some point, that there [a place indexed via perception] is the 8:25 train. Similarly, should my life be endangered by an approaching bear, it might help me to know it. But it will not be enough for me to know of this danger to me under some impersonal description of me such as “the person sitting in Bruno’s favorite berry patch” or even under the name “Ruth Millikan” unless I further know that I am the person in Bruno’s favorite berry patch or that I am Ruth Millikan (I might not know, for example, should I be amnesiac). But this kind of thought—there is the 8:25, I am Ruth Millikan—is, it is supposed, indexical. (1990, p. 723, bolded emphasis added)

Here is Colin McGinn on being-in-the world, agency, and indexicality:

But there is also the kind of thought we operate with when actively engaged in the world, when the world is directly presented to our consciousness; and this seems shot through with indexicality.... It thus seems that what the phenomenologists would call 'being-in-the-world' presupposes indexical thought. (1983, p. 91)

The present suggestion, then, is that indexical concepts are ineliminable because without them agency would be impossible: when I imagine myself divested of indexical thoughts, employing only centreless mental representations, I eo ipso imagine myself deprived of the power to act. (1983, p. 104)

Here is David Chalmers on indexicality and “the world as I find it”:

The indexical fact expresses something very salient about the world as I find it: that David Chalmers is me. How could one explain this seemingly brute fact? Indeed, is there really a fact here to be explained, as opposed to a tautology? The issue is extraordinarily difficult to get a grip on, but it seems to me that even if the indexical is not an objective fact about the world, it is a fact about the world as I find it, and it is the world as I find it that needs explanation. The nature of the brute indexical is quite obscure, though, and it is most unclear how one might explain it. (1996, p. 85)
Here is John McDowell on the first person and consciousness:

Continuous “consciousness” is intelligible (even “from within”) only as a subjective angle on something that has more to it than the subjective angle reveals, namely the career of an objective continuant with which the subject of the continuous “consciousness” identifies itself. The subjective angle does not contain within itself any analogue to keeping track of something, but its content can nevertheless intelligibly involve a stable continuing reference, of a first-person kind... (1998, p. 303)

Here is David Velleman on Kant, agency, and “the first person”:

Kant’s framing his maxim in the first person is no accident. He could not have restated it, for example, as “Immanuel Kant will make lying promises when he is in need.” Such a third-personal thought would not be a maxim of action, since it could not be acted upon by the thinker until he reformulated it reflexively, in the first person. Insofar as the target of universalization is a practical thought, it is essentially first-personal. (2006, p. 121)

Here is François Recanati on the widespread acknowledgment of irreducible indexical content:

No transformation from indexical to non-indexical is possible without affecting the cognitive significance of the utterance and therefore changing the thought it expresses....The irreducibility and indispensability of indexicals is widely acknowledged. (2007b, p. 243)

Here is Susanna Siegel on the significance of indexicality in perceptual content:

Many contentful experiences seem to have contents that must be specified by the use of indexical expressions, such as “over there”, “to the left”, “here”, “in front of/behind me”, “just a second ago”, “since a few second ago”, and so on. For instance, some auditory experiences seem to present sounds as coming from a direction relative to the ears; some proprioceptive experiences seem to present pressure sensations as located in specific parts of the body; visual experiences seem to present things in locations relative to the eyes and the rest of the body. More generally, contentful experiences present the world from the subject’s perspective both in space and in time. (2011)

Here is Héctor-Neri Castañeda on powers of deliberation and indexical thought:
When I act intentionally on an object, I care neither about the Russelian (Kaplanian) propositions beyond as such nor about their Fregean (Kaplanian) first referents. These are not crucial. What I need is to bring those doxastic referents somehow into my experience through my thinking indexical (that is, experiential) references...the singular proposition that mobilizes my powers of deliberation and action is the indexical truth within my grasp. (1989, p. 126)

We could go on and on, but we end this section with a passage from Castañeda that maybe summarizes much of what came before: “Indexical reference is personal, ephemeral, confrontational, and executive. Hence, it is not reducible to nonindexical reference to what is not confronted” (1989, p. 70).

We hope these passages substantiate our claim that the Perry–Lewis line has had immense impact throughout philosophy. It isn’t much of an exaggeration to say that their view has a unique status in contemporary philosophy: it is widely relied on without argumentation and there is hardly any critical examination of it. This is a dangerous combination, and it is in such an intellectual context that we think our book can serve a purpose. We hope to generate in our readers at least a bit of doubt about the foundational assumptions in this domain.

1.4. A More Systematic Presentation of Alleged Repercussions of Essential Indexicality Throughout Philosophy

Section 1.3 used quotes to document the widespread endorsement of what contemporary philosophers take to be Perry and Lewis’s view. This section is devoted to a more systematic presentation of the assumed repercussions of a theory of indexicality in philosophy. These obviously begin in philosophy of language, and in the role of a theory of content in the philosophy of mind. An account of how indexicals function is central to any attempt to understand natural languages, communication, and representation more generally (both linguistic and mental, and the mental which includes both the doxastic and the perceptual). Pre-theoretically, indexicals appear to be devices that put us in a uniquely direct and primitive kind of contact with
ourselves ("I"), with important features of our environment (the time and place we are at, as with "now" and "here"), and with objects we can demonstrate ("that"). Crafting an adequate semantic account of that direct and primitive contact has proved a persistent challenge. We take Kaplan’s account of indexicality to be the most philosophically minimal and conservative account since it preserves a level of content at which there is no difference between indexical and non-indexical expressions. Even this conservative approach emerges only after a great deal of work in the semantics of indexicality. And numerous less conservative options compete with the Kaplanian approach. At the origin of analytic philosophy we find Frege struggling to account for first-person and demonstrative thought in a framework that distinguishes sense from reference and assumes that senses are public. Reflections on “I” led Frege to say slightly mysterious things like:

Now everyone is presented to himself in a particular and primitive way, in which he is presented to no one else. So, when Dr. Lauben thinks that he has been wounded, he will probably take as a basis this primitive way in which he is presented to himself. And only Dr. Lauben himself can grasp thoughts determined in this way. (1956, p. 298)

Just how this is compatible with a framework in which senses are public is a challenging question for those engaged in Frege exegesis and we will not attempt to answer it in this work. Putting aside exegetical issues, it is often assumed that indexicals raise particular challenges for defending the kind of descriptivism often associated with Frege and his followers. Attempts to answer these challenges motivate important work by Evans (1982), McDowell (1977), Forbes (1987), Peacocke (1981), Kripke (1980), and others. We address Fregean approaches to indexicality, and the tensions between indexicality and descriptivism, in more detail in Chapter 4.

Another important tradition in the semantics of indexicality derives from Lewis 1979. In that paper Lewis argues that we should give up the traditional view that propositions are the semantic values of sentences and the objects of attitudes such as belief, hope, and fear. Lewis proposes that the fundamental semantic coins should be not propositions but properties, and that talk of agents standing in relations to propositions should be replaced with talk of what he calls “self-attribution”
of properties. Lewis’s seminal work has given rise to two important traditions in the theory of indexical content. The use of centered worlds as semantic values and objects of attitudes has become increasingly widespread among semanticists and philosophers of language, and the centered-worlds framework has played a central role in the disputes between contextualists and relativists. Centered worlds play a foundational role in the variety of semantically motivated relativism found most prominently in the work of Andy Egan (see, inter alia, his 2006a, 2007, 2009, and 2011). On Lewis’s semantic treatment, the properties self-attributed by the agent are essentially perspectival contents without involving representation of the perspective because the self-attributing agent is not represented in the property expressed by sentences and beliefs. This aspect of Lewis’s work has also motivated the project of primitive unarticulated representation in the work of François Recanati (2007b), which centrally features implicitly de se states which are about the self without deploying a self-representation.

The Fregean and Lewisian traditions loom especially large in the literature on the semantics of indexicality, but other approaches also attempt to account for what are taken to be special problems of indexicality using resources less conservative than the minimal Kaplanian framework. For example, Higginbotham (2003, 2010) uses self-reflexive contents and Kamp (1971, 1981, 2001, 2007) uses discourse representation theory.

From these starting points in the philosophy of language, the ramifications of issues about indexicality radiate outward. Concerns about the distinction between indexical de se mental states and non-indexical de dicto or de re states play a central role in the philosophy of action (see classics such as Hare 1971, Kenny 1973, Castañeda 1975, and more recently Owens 2011). Classic examples from Castañeda and Perry are often taken to suggest that no action is possible, or rationalizable, without the presence of some indexical-containing belief, which serves to tie “objective” reasons for action with the particular embodied situation of the acting agent. For many of our contemporaries, we expect

---

5 Many of the issues at the center of these debates are also found in discussions surrounding temporalism and eternalism about propositions. See Prior 1959, Richard 1981, Ludlow 1999, and Brogaard 2012.
these considerations—inspired in large part by the classical examples
construed by Perry—provide the most powerful and compelling rea-
sons for endorsing the kinds of views we oppose in this work. We
suspect that many of the other motivations are, at least in part, parasitic
on the consideration that appears to give indexical thought a central
role in an understanding of rational agency.

Indexicality also plays a role in the literature on the philosophy of
perception (see Egan 2006b, Brogaard 2011, Chalmers 2006, and Lewis
1979). Because our perceptual states typically represent the world as
seen from a perspective, it is sometimes thought that some distinctively
indexical kind of content is needed to characterize those states. The per-
ceptual sameness between the states of two agents each of whom sees
a tree to their left is taken to call for a content-level sameness, with the
two agents both having perceptions with the indexical content “tree to
my left.” Indexical contents are also used to approach the role of percep-
tion in yielding secondary qualities, which in some sense characterize the
world not as it objectively is but as it is subjectively given to us (Egan
2006b). The centered-worlds framework has recently been influential
here, allowing, for example, an account of color perception according to
which phenomenal red is a perspectival, self-attributed property.

Questions of indexical content also impact issues in metaphys-
ics. The semantics of “I” is often thought to be relevant to personal
identity and self-consciousness (see, for instance, Shoemaker 1963,
1968/1984, 1994; McGinn 1983; Chalmers 1996; McDowell 1998; and
Nagel 2012, among many others). Some think the semantics of “now”
and tense more generally has implications for the metaphysics of time.
Recent work by Yalcin and Lee (Lee 2006, Lee and Yalcin 2010) on
the objectivity of left–right properties interacts with the semantics of
indexicality, and raises general issues about the proper metaphysical
characterization of the state of the world by those who take many
properties to be instantiated only relative to a perspective. The meta-
physical issues have further ramified into issues in the philosophy of
religion in Mark Johnston’s recent work both on the relation between
personal identity and the possibility of surviving death, and on the
connection between the roles of de se thought in motivation to action
and the proper understanding of agape (2010, ch. 3; 2009, ch. 6). Other
issues in the philosophy of religion connect with de se contents—it has
been argued that divine omniscience is incompatible with the existence of proprietary first-person knowledge (see Kretzmann 1966 and Grim 1985). According to, e.g. David Velleman and Caspar Hare, an appeal to \textit{de se} attitudes play an important role in understanding central issues in moral philosophy, such as moral relativism (see Velleman 2013, Kant’s Maxim (see Velleman 2006), and the best formulation of egocentric hedonism (see Hare 2007).\footnote{Hare (2007, p. 372) asks us to perform the following experiment:

Today, many hundreds, if not thousands, of Russians will spill boiling water on their hands. Pour boiling water on your own hand and compare your present discomfort with the absent discomfort of the northernmost Russian spiller. Which is worse?

He suggests that our reaction will be “This pain is dreadful, worse than any of those Russian pains,” and goes on to argue that fitting this reaction together with a recognition of the symmetry between our situation and that of the northernmost Russian spiller requires \textit{perspectival} facts about the presentness of pains.}

Issues in epistemology also turn on the proper treatment of indexicality. The alleged phenomenon of immunity to error through misidentification has a long and distinguished history in epistemology. According to one dominant trend, it has its source in first-person (and maybe other indexical) thoughts and attitudes (see Christofidou 1995, Smith 2006, Hamilton 2007, and O’Brien 2007). A central question in recent work in formal epistemology has been the proper dynamics for updating probabilistically characterized belief states in light of indexical information (the “Sleeping Beauty” problem), with different formal frameworks relying on different characterizations of indexical content and different assumptions about the relation between that content and non-indexical content (Elga 2000, Arntzenius 2003, Meacham 2008, Pust 2012, Titelbaum 2013).\footnote{We have occasionally been told in conversation that some of the crucial arguments in favor of essentially \textit{de se} contents emerge from considerations of the Sleeping Beauty problem. However, we find ourselves in agreement with Titelbaum’s diagnosis that the Sleeping Beauty problem has no essential connection with the theory of content and that formal updating models can be created to handle it which are indifferent to questions of underlying content. We won’t rehearse Titelbaum’s arguments here—for detailed discussion, see Titelbaum 2013 (especially §9.2) and his (forthcoming). We remain open to being convinced that the motivation toward the \textit{de se} here is deeper than we have realized, but given our current diagnosis, we do not discuss Sleeping Beauty problems further in this book.}

These are but a few salient examples. We will not pursue all of these topics in what follows. Our primary focus will be on the foundational issues concerning indexical content and we will address these various
applications only insofar as they contribute directly to resolving the foundational issues.

1.5. More Precise Articulation of Our Target: Essential Indexicality Theses and the Alleged Explanatory Roles of Indexicality

A lot of what follows will be an attempt to make precise just what we are opposing. As is usual in philosophy, familiar labels and slogans turn out to be obscure and unhelpful for serious work. A theme that runs through this work is that the cluster of theses often thought to concern “essential indexicality” can be usefully broken down into a number of sub-theses, each of which concerns various explanatory roles indexical language and thought is alleged to have, e.g. in semantics, philosophy of mind, philosophy of action, theory of perception, and epistemology. Our focus in this work will be on these various sub-theses. There are many candidate explanatory tasks and they are interconnected in various ways. As a result, attacking Essential Indexicality is a difficult task, because a large collection of motivations and concerns come together in a kind of “teepee” argument, in which a number of argumentative planks intersupport each other. No plank is sufficiently strong to establish the position, but each lends credibility to the others because the cluster of considerations gives the appearance of a unified phenomenon. A thorough debunking thus requires separating the planks. That’s what we aim to do in this work. Our focus will be on six kinds of alleged explanatory roles:

**EI: Agency.** Chapter 3 explores the idea that indexicality (and “the de se” in particular) plays an essential role in explaining and rationalizing action.

**EI: Opacity.** Chapter 4 addresses the question of whether the presence of indexicals in (apparently⁸) opaque contexts raises

---

⁸ We will from now on omit the word “apparent” when we talk of opacity and substitution failures. It is our view that opacity is just an appearance—co-referential terms can always be replaced without a change in truth value. We will not argue for that position here and
questions that are fundamentally and interestingly different from general issues about opacity.

**EI: Information.** Chapters 5 and 6 evaluate the Lewisian thought that indexicality plays an essential role in explaining information gathering.

**EI: Epistemic.** Chapter 7 concerns the question of whether immunity to error through misidentification is an essential feature of indexicality (if it is a feature of anything at all).

**EI: Perceptual Content.** Chapter 8 explores the line of thought that if there is such a thing as perceptual content, it has to encode the perspective from which the perceptual experience is taking place in an indexical manner.

**EI: PRO.** Chapter 9 explores the view that a distinctive kind of indexical content is essential for the interpretation of PRO constructions in natural language (as in sentences like “John remembers winning the race”).

We should highlight right away that “indexicality” as it is used in articulating these various theses isn’t yet well defined. In leaving the interpretation of that expression open at this stage we follow Perry (and the tradition following his lead): we want to leave open what level (or kind) of content the explanatory role is assigned. However, our working hypothesis is that no matter how “indexicality” is spelled out, all of these theses are false.

### 1.6. Basic Kaplanianism and the Goals of This Book

According to what we will label “Basic Kaplanianism,” each of the sentences “Smith is happy” and the sentence “I am happy,” as uttered by Smith, has two levels of meaning. Each has a *content* and a *character*.

- The *content* of each sentence (in context) is its truth-conditional representation. This can be thought of as a Russellian structured
proposition or as a set of possible worlds; these finer distinctions won’t matter for most of our purposes.

- The character of each sentence is then a function from contexts to contents expressed by the sentence in context. Indexicals are characterized as having non-constant functions. This is just a slightly theoretical way of saying that they change referents between contexts of utterance.

“Smith is happy” and “I am happy” have different characters. The character of “Smith is happy” is a constant function that maps every context to the same content. The character of “I am happy,” on the other hand, is a non-constant function that maps each context to (for example) the structured proposition pairing the speaker in that context with the property of being happy. Moreover, it is no accident that the two sentences have different characters. Trivially, no non-indexical sentence can have the same character as an indexical sentence.

The basic Kaplanian framework makes use of a very coarse-grained notion of content: contents as functions from worlds to truth values. But a central insight of the framework—the isolation of context sensitivity into a character level of meaning—can be preserved even if one prefers a different theory of content. Characters can just as well be functions from contexts to structured propositions, or from contexts to Fregean thoughts. What persists across such variations is that the implementation of context sensitivity resides entirely in character, not in content.

Here’s a way to articulate some of the basic motivations behind this book:

*We hope to at least move our readers towards the view that all there is to context sensitivity is the undeniable fact that some expressions have a non-constant character. Once the characters (constant and non-constant) of expressions have been characterized, there is no further need for additional devices like “first-person concepts” or “demonstrative concepts.” Moreover, the considerations coming out of the Perry–Lewis tradition give us no reason to change our theory of content, and provide no evidence that there are philosophically interesting or important roles played by non-constant characters.*

Note that these rough claims includes important weasel words such as “interesting” and “important.” The weasel words are crucial,
because there’s clearly some role played by the indexical, i.e. non-constant, characters. It’s the specific non-constant character of “I,” for example, that explains why Smith says “I am happy” to Jones in order to express Smith’s happiness, rather than “You are happy.” The claim is just that such roles aren’t interesting, or philosophically deep and illuminating. We don’t have a general theory of the philosophically deep and illuminating, so we don’t have a direct way of proving that indexicality isn’t linked to the philosophically deep and illuminating. So we will approach the issue piecemeal, by considering various proposals (i.e. the various EI-theses) for what is deep and illuminating about indexicality, and arguing that those proposals systematically fail.

1.7. Some of the Many Issues We Will Not Address

There are many interesting and controversial questions about indexicals and in particular their semantics that we will try to steer clear of in what follows. Our working assumption is that the specific issues about Essential Indexicality that we focus on here are largely independent of whether there are monsters in Kaplan’s sense (see Kaplan 1989 and Schlenker 2003), the role of intentions in indexical reference fixing (see Castañeda 1992 and Akerman 2010), the interpretation of indexicals used on billboards and answering machines (see Egan 2009), the interpretation of indexicals in mixed quotation (see Maier 2007), and whether, e.g. “knows” or “might” is context-sensitive (see, e.g. Kratzer 1977, Hawthorne 2004, and Cohen 1986). These are but

9 The relation between (a) what is potentially “important” or “interesting” about characters and (b) the general issue of opacity and substitution failures, is a complicated one, and will not become fully clear until the discussion of opacity in Chapter 4. We do not mean to deny that differences in character can give rise to differences in substitution behavior. What we do deny is that character as such is what is philosophically interesting here. What is philosophically interesting is the broader category of Fregean sense (or whatever you think is the right explanation for opacity), which encompasses both indexical and non-indexical cases. Difference in character is then just one way (and, we argue in Chapter 5, not an interestingly distinctive way) to have difference in sense.
a few illustrations of the many important research questions about indexicals that won’t be discussed here. We mention them simply to re-emphasize that our concern is specifically with the various kinds of Essential Indexicality claims outlined earlier. It goes far beyond the scope of this book to present a full theory of all of the interesting properties of indexicals.

1.8. How to Read This Book

Here is a brief introduction to the various chapters and their interconnections.

• Chapter 2 briefly addresses three background issues that will come up repeatedly throughout the book: (i) whether Essential Indexicality is about mind or language, (ii) whether some indexicals (e.g. “I”) are philosophically more significant than others, and (iii) the connections between arguments for opacity and arguments for Essential Indexicality.

• We suspect that Chapter 3 is the most important chapter in this book. In that chapter we argue that there are no interesting or distinctive explanatory connections between indexicality and agency. In talking to philosophers about these issues over many years, we find that attempts to articulate the essence of Perry and Lewis’s insight tend to come back to something or other about a distinctive connection to agency. We also find that when other kinds of Essential Indexicality are appealed to, they are often seen as parasitic on some kind of agency connection.

• We suspect that Chapter 4 is the second most important chapter in the book. It concerns the role of indexicality in an account of opacity. The topic of this chapter is elusive (since opacity is elusive), but roughly we make four points: (i) the claim that co-referential indexicals can’t be substituted in attitude (and other allegedly opaque) contexts isn’t surprising and can’t be what the thesis of Essential Indexicality amounts to, (ii) there’s no evidence of a distinctive kind of opacity for indexicals, (iii) Frege had a problem accounting for indexicals, but that just shows how weird
Frege’s theory was—it shows nothing intrinsically interesting about indexicality. (iv) at the end of this chapter we explain how the goals of this book can be achieved even though we do not argue for any particular theory of opacity (or, indeed, any particular theory of content).

• We suspect that Chapter 5 will be considered the most important chapter by Lewis aficionados and those whose work relies on the proposals in Lewis’s paper “Attitudes De Dicto and De Se.” We show two things: (i) that Lewis provides no new evidence for the existence of (or need to recognize) anything irreducibly de se; (ii) that even if you are unconvinced by our arguments in this book and persist in thinking that there are irreducibly de se attitudes, Lewis’s theory doesn’t explain or illuminate what those are.

• Chapter 6 considers a reply on behalf of Lewis: that appeal to his functionalism avoids the objections we raise in Chapter 5. This chapter will be of interest primarily to a subset of Lewisians and to those non–Lewisians who think that some version of functionalism constitutes an argument for Essential Indexicality.

• In Chapter 7 we turn to issues in epistemology. It is often assumed that indexical contents have an intimate and distinctive connection to the alleged phenomenon of immunity to error through misidentification. One might try to anchor Essential Indexicality in such connections. We argue a) that there is no such thing as immunity to error through misidentification, and b) even if you don’t endorse our arguments in a), the phenomenon has no distinctive connection to indexicality.

• Chapter 8 explains why indexicality does not play an essential or irreducible role in an account of the contents of perception. Surprisingly, there are few contemporary philosophers of perception who explicitly argue that it does, so we don’t have a large group of explicit opponents here. That said, we think we should have opponents: if de se content really plays an essential or irreducible role in an account of content, it would be extremely surprising if it didn’t also play an important role in an account of perceptual content. So we think the fact that it clearly doesn’t play such a role is an additional argument for our deflationary position.
Chapter 9 concerns the question of whether the semantics of PRO constructions can be helpful to those who think there is a distinctive kind of \textit{de se} content. We show that it is not helpful or illuminating. This chapter is primarily of interest to those inclined to think PRO constructions could help illuminate or explain irreducibly \textit{de se} attitudes. We suspect this chapter might have the narrowest audience of all the chapters.

Deflationary projects such as ours are in danger of seeming non-constructive. This is wrong—both in general and in this particular case. A deflator of some phenomenon, $F$, will often start out by diagnosing mistakes made by $F$-theorists. However, when that is done, the deflator is left with an $F$-less world and faces the task of explaining all the allegedly $F$-involving phenomena with a lighter toolbox. This is often difficult constructive work.

Throughout this book we show how various specific explanatory tasks can be done without any essentially indexical properties. In Chapter 10 we broaden our perspective and sketch a positive picture of a world and our role in it without anything essentially perspectival.

\[\text{10} \text{ Though not always; sometimes deflation just removes an illusion and all that’s left is explaining why so many people were under the illusion.} \]
Chapter 2

Preliminaries:
Language–Mind, Super Indexicals, and Opacity

This brief chapter deals with three preliminary sets of issues, which are important for thinking about Essential Indexicality and will come up repeatedly throughout our discussion of the various Essential Indexicality theses in later chapters. It is, we think, helpful to highlight them at this early stage.

2.1. Essential Indexicality:
Language or Mind?

These issues have traditionally been approached from two closely connected fields: the philosophy of language and the philosophy of mind. Within the Perry–Lewis tradition, some considerations derive primarily from the philosophy of language. The focus is then on linguistic indexicals—linguistic devices such as “I,” “now,” and “that” whose contribution to content depends on the context in which they are used. A central topic of investigation is the nature of the contents we express using indexicals and context-sensitive terms more generally. Parallel to these linguistic considerations, there is an approach that focuses on mental or conceptual indexicality (or indexicality “in thought”), as in some of Lewis’s work and some of Perry’s later work. This approach aims to describe indexicality (and perspective) not as it attaches to specific linguistic expressions, but as it characterizes non-linguistic cognitive states. While this approach often is parasitic on the linguistic
approach in the sense that it sees the cognitive indexicality as derived from, or at least modeled on, linguistic indexicality, this is not necessarily so. In what follows, we will move back and forth between these two approaches. When the literature we are engaging with focuses primarily on the language-first approach to indexicality, this will also be our approach to these issues, but whenever needed, we will switch perspective and focus instead on various proposals for how to understand indexicality at the level of thought. Our conclusion is the same with respect to both approaches: the radical claims made about indexicality are unsubstantiated and fundamentally mischaracterize the nature of indexicality, content, perspective, and the connections between them.

These quick moves back and forth between language and mind are not unproblematic and since we will largely ignore those problems later, we will quickly sketch some of them here. A famous example from John Perry provides a good illustration. Perry says:

Let us imagine David Hume, alone in his study, on a particular afternoon in 1775, thinking to himself, “I wrote the Treatise.” Can anyone else apprehend the thought he apprehended by thinking this? (1977, p. 487)

This passage illustrates a common and problematic way of presenting these kinds of issues. A sentence is quoted, and then we are told that Hume thinks that sentence. That kind of claim is very confusing. First, it doesn’t mean that the agent thinks the quoted sentence—i.e. thinks about the quotation. It is supposed to mean something like: the agent thinks the thought expressed by that sentence. But not quite, since we should all agree that in some sense there are many ways to think, e.g. the thought that Hume wrote the Treatise. That thought can, arguably, be expressed by both the sentences “Hume wrote the Treatise” and “I wrote the Treatise” (when uttered by Hume). And so, the more charitable interpretation might be something along the lines of: Hume thought what he would have expressed by an utterance of “I wrote the Treatise” and he did it using a mental analogue (or correlate) of the public-language sentence, “I wrote the Treatise.” Something like this interpretation is supported by the continuation of the passage:

First note that what he thinks is true. So no one could apprehend the same thought, unless they apprehended a true thought. Now suppose Heimson is a bit crazy, and thinks himself to be David Hume. Alone in his study, he says
to himself, “I wrote the Treatise”. However much his inner life may, at that moment, resemble Hume’s on that afternoon in 1775, the fact remains: Hume was right, Heimson is wrong. **Heimson cannot think the very same thought to himself that Hume thought to himself, by using the very same sentence.** (1977, p. 487, bolded emphasis added)

Perry here seems to assume that we use the public-language sentences when we think the thoughts in question. This assumption is probably just a kind of heuristic: it would be uncharitable to take it as a serious theoretical commitment. Nonetheless, this kind of move between claims about language and mind the passage illustrates is problematic. While it is relatively clear what context sensitivity amounts to when we talk about linguistic expressions, it is less clear what indexicality amounts to at the level of belief and other propositional attitudes. Often this issue is skirted in the literature by the introduction of terms like “de se attitudes” where it is left unsaid what, if anything, such attitudes have to do with the first-person pronoun. A full story about the nature of mental indexicality needs to clarify at least the following two issues.

First, we need to figure out which beliefs are going to count as “indexical” beliefs.¹ In the linguistic case, it looks like the idea of indexicality gets introduced in the following way. We want a general procedure for associating linguistic entities with contents. But on examination, we discover that there can be two linguistic acts that have the same syntactic/lexical form, but have different content. (For example, take two utterances of “I am a philosopher” that differ in truth value, despite being composed of the same words arranged in the same way.) In order to account for the combination of sameness and difference, we identify some portion of the utterances as indexical. Then we can have sameness of character and difference of content for the two occurrences of the indexical expression. But applying this model to the mental case isn’t straightforward. In order to motivate introducing indexicality, we need to have two mental states that display the same combination of sameness and difference. The difference part is easy: we want the two mental states to have different contents (as

¹ Our focus here is on belief, but the points we make are meant to apply to all propositional attitudes.
marked by truth conditions, inferential role, cognitive role, etc.). But the sameness is trickier. In the linguistic case, we could appeal to same-ness of syntactic/lexical form. Now, maybe we can do the same thing in the mental case. If we hold a strong “language-of-thought” thesis, then there will be something that counts as the “linguistic form” of a thought to which we can appeal.\footnote{Although the version of language of thought would need to be quite strong. For example, in order to get a sensible story about indexicality matching that of the personal pronouns, we need our two mental states to occur in different individuals, which means we need our language of thought to be interpersonally robust, in the sense that we can identify the same linguistic form across multiple agents. A description of the mental state at a neurologi-cal or computational level, for example, would be unlikely to be robust enough.} Without the language-of-thought thesis, it’s harder to see how we are going to identify the “same belief” in the relevant sense. Maybe the best option is to rely on “sameness of functional role.” But then we’d better not be functionalists about intentional content (for further discussion of the role of functionalism in these debates, see Chapter 6).

Second, we need to figure out what counts as the “context” of a mental state. We won’t dwell on this worry, but it is a real worry. Some aspects of the linguistic notion of context may transport easily enough: “speaker” looks like it may transition seamlessly to “thinker.” But “audience” looks like a much more complicated notion. One’s thoughts aren’t by virtue of the physical setting directed towards anyone, but nevertheless it looks like one can have “you”-thoughts. And time is very tricky. Since beliefs can be non-occurrent and ongoing, it’s not at all clear that we can identify anything helpful that counts as the “time of the thinking”—perhaps the full interval in which the belief is held?

Some philosophers in the Perry–Lewis tradition are sensitive to these issues. Building on work by Evans, Perry, and Bach, Recanati attempts to show that the indexical model can be extended to the mental in recent work. It will be useful to briefly introduce some of these ideas. Recanati says: “The linguistic phenomenon of indexicality is relatively well understood, so it is worth enquiring whether the indexical model applies in the mental realm” (2013, p. 57). His proposal

\footnote{One worry with this strategy: it will be hard to make the case that “I” is a context-sensitive term, i.e. that it has a non-constant character, since no single agent is ever aware of that supposed mental indexical ever changing extension.}
is that a theory of mental files can be used to characterize something like mental indexicality. He says:

Mental files, I will argue, possess the essential features of indexicals....Obviously, the notion of conventional meaning does not apply in the mental realm. But at least the type/token distinction applies. As far as mental files are concerned, they are typed according to the type of ER [epistemically rewarding] relation they exploit. Thus the self file exploits the relation in virtue of which one can gain information about oneself in a way in which one can gain information about no one else (as Frege puts it). My self file is not the same as yours, and they refer to different persons, of course, but they belong to the same type: they are both self files, unified by the common ER relation it is their function to exploit. We see that the function of files—namely, informational exploitation of the relevant ER relation—plays the same role as the conventional meaning of indexicals: through their functional role, mental file types map to types of ER relations, just as, through their linguistic meaning (their character), indexical types map to types of contextual relation between token and referent. The indexical model therefore applies to mental files, modulo the substitution of functional role for linguistic meaning. (2013, pp. 57–60)

Recanati’s proposal illustrates the kind of functionalist move gestured at earlier and this is not the place for a thorough evaluation of that move. We include it here simply to illustrate how one might try to extend the notion of indexicality to the mental and how doing so is far from trivial and comes with very heavy theoretical commitments. More generally, insofar as the literature tends to move between talk of indexicality (and the de se) in language and in thought, we are owed—but seldom given—a theory of how to understand those transitions.

One option should be mentioned here: our opponents could argue that the label “essential indexicality” was poorly chosen and that what’s important about the Perry–Lewis tradition is just that indexicals denote special “ways of thinking of objects.” Reflection on indexicality in language is a heuristic for getting us to pay attention to those distinctive states. It is not important, according to these views, that those states themselves are indexical. So the kinds of issues in this section are not pressing. We should emphasize that this kind of position is not a primary target in this work since it is compatible with our rejection of any kind of distinctive role of indexical content. That said, we think
it is false as well and we will point out why as we go along (in effect, we take the discussion in Chapter 6 to be a refutation of this position).

2.2. Are All Indexicals Equal or Are There Super Indexicals?

Proponents of Essential Indexicality face an important choice point: are all indexicals equally essential or is there a proper subset of them that are more essential than the others—call them Super Indexicals? Those endorsing the latter view often hold that it derives from the special connection between the first person and agency. Here is a passage from Colin McGinn expressing that attitude:

...all the [essential] indexicals are linked with I, and the I mode of presentation is subjective in character because it comprises the special perspective a person has on himself. Very roughly, we can say that to think of something indexically is to think of it in relation to me, as I am presented to myself in self-consciousness. (1983, p. 17)

On this kind of view, first-person content is distinctive (or essential) and if other kinds of content have distinctiveness (or are essential), it is inherited from their relationship to first-person content. The privileging of the first person has a long and distinguished philosophical history, tracing back at least to Descartes and Kant, and strong echoes of that tradition can be found in David Lewis’s mindboggling claim that every thought we have (and every claim we make) is a thought (claim) about ourselves (1979, p. 522). A view that designates “I” as the Super Indexical face challenges: it needs to show that whatever distinctive properties are possessed by “I” are not also found in connection with, e.g. “now,” “here,” and “you” (otherwise “I” would not be special). While many of the theorists we discuss in what follows pick the first person as their Super Indexical, it is not the only option. Kaplan (1989) discusses what he calls “the demonstrative theory of indexicals.” While Kaplan attributes this view to “the sloppy thinker,” we’re not convinced that only sloppy thinkers would be attracted to it (p. 534). According to it, all indexicals are tacit demonstratives: “I” means the same as “this person,” “now” the same as “this time,” “here”
the same as “this place.” In this way, all indexicality is a form of demonstration and understanding demonstratives becomes the key to an understanding of (at least the core) indexicals.

Proponents of Essential Indexicality don’t have to designate a Super Indexical, but doing so might help them provide an explanation for Essential Indexicality. If a distinctive feature reliably tracks indexicality, it would be nice to know why it’s exactly the indexicality that matters. One kind of answer is a direct theoretical answer: one shows how indexicality in itself explains the status of the Super Indexical in a way that non-indexicality cannot. Something like this is maybe available in some cases, but often this kind of direct argument is lacking. For example, it’s not at all clear why indexicality per se should be crucial to explanation of action. So a different sort of strategy is to suggest that the Super Indexical originates not in indexicality per se, but in some specific indexical construction (perhaps the indexical self-conception (as suggested by McGinn) or the nature of demonstratives (as suggested by Kaplan’s sloppy thinker). Maybe it then extends to other indexicals if one can specify other indexicals in terms of the Super Indexical.

It is worth marking some further options in this connection: one could hold the view that there’s more than one Super Indexical. Maybe the first person has distinctive features not shared with “now” and “here” and “that.” Those three might again each have features not shared by the others. In general, one should not assume from the outset that Essential Indexicality is a unified phenomenon across indexicals. Maybe “now” has an important distinctive role in action explanation that “here” does not have and maybe first-person thoughts play a role in perception that “now”-thoughts don’t play. We won’t take a stand on any of these issues now. The point is simply to emphasize that it is too often assumed that claims about “essential indexicality” apply across the board to all indexicals. An equally plausible working assumption is that the various interesting Essential Indexicality theses will apply more locally to specific subsets of indexicals.

Finally, it is worth noting that some of the most heated debates in the philosophy of language in the last years have been over the scope of indexicality (and context sensitivity more generally). An enormous amount has been written about whether and how, e.g. “all,” “know,”
“might,” “delicious,” “good,” “if, then,” “and,” “red,” “just,” “justified,” “probable,” “local,” “ready,” and “left–right” are context-sensitive. Those who think all context-sensitive expressions are distinctive along some dimension should have a view about which expressions count as context-sensitive. The more widespread one takes the phenomenon to be, the more pressing the issue in this section becomes: are all these expressions distinctive or just a subset? If just a subset, what is distinctive about that set and why doesn’t it extend to all context sensitivity?

### 2.3. Opacity and Essential Indexicality

It is a familiar and much investigated fact that the principle of substitutivity of co-referential terms appears to fail in certain kinds of linguistic contexts. A recurrent theme in this work will be that much of what has been paraded as instances of essential indexicality is nothing but an instantiation of this (apparent) phenomenon. It should come as no surprise that indexicals behave like, e.g. proper names in this respect. Two methodological points are important to highlight right away (and these will be pursued further in Chapter 4).

First, Essential Indexicality more or less entails opacity. Suppose that indexicals create a special kind of content that can’t be created by non-indexicals. Now suppose we can talk about that content—at least in the minimal sense that there’s some linguistic setting that’s sensitive to the distinction between that content and others. Without such a setting, it’s hard to see how we could be convinced that there’s a special kind of content. Then in the context of that linguistic setting, we can’t replace indexicals with co-referring non-indexicals, so we have opacity. Indeed, it looks like some form of opacity will always be needed as evidence for Essential Indexicality.

Second, it now becomes crucial to note that if a linguistic setting $S$ is opaque for all singular terms (so that, for example, it doesn’t allow substitution of “Clark Kent” for “Superman” *salva veritate*), then it’s going to be less helpful as evidence for Essential Indexicality. It’s not that it lends no support, but there’s always the following worry: when we discover that we can’t replace, e.g. “I” with a co-referring singular term, we don’t really know if that’s because of the Essentiality of
the indexicals, or just because of the general opacity. So what would be ideal for establishing Essential Indexicality would be linguistic settings that are opaque along the indexical/non-indexical line, but not opaque for singular terms in general. One of our strategies to support the claim that there are no such settings will be to try to show that a variety of linguistic settings lack the required feature.
Chapter 3

Indexicality, the *De Se*, and Agency

Is there a philosophically significant connection between indexicality and agency? According to orthodoxy the answer is a resounding “yes.” It is more or less common ground in contemporary philosophy that some such significant connection exists. It is taken for granted to such an extent that it by now seems unnecessary to provide arguments for it—all one needs to do is include a footnote referring to Perry 1979 (or Lewis 1979) and that's supposed to settle it. Consider the following passages (some of which were quoted above):

The present suggestion, then, is that indexical concepts are ineliminable because without them agency would be impossible: when I imagine myself divested of indexical thoughts, employing only centreless mental representations, I *eo ipso* imagine myself deprived of the power to act. (McGinn 1983, p. 104, bolded emphasis added)

It is widely agreed that agents need information in an egocentric form: they must think of places as “here” and “there”, times as “now” and “then” if they are to be able to act on what they know (Perry 1979). (Owens 2011, p. 267)

A strong contemporary current runs to the effect that the ability of an agent to project knowledge of the world into relevant action in the world depends upon the ability to think indexical thoughts. (Millikan 1990, p. 723, bolded emphasis added)

...practical guidance is, in Perry’s phrase, essentially indexical, in the sense that its function depends not only on which of many propositions it expresses but

---

1 Or egocentric thought, or the first person, or the *de se* or (insert your favorite piece of relevant terminology).
also on how that proposition is determined by the context—specifically, on its being determined in the same way as the reference of indexical expressions such as “I”, “you”, “here”, and “now”. (Velleman 2013)

This chapter has two goals. First, to try to make precise exactly what connections are supposed to obtain between indexicality and agency. Second, to evaluate arguments that can be given for those alleged connections. Two points are worth highlighting right away:

a) It is probably uncharitable to read Perry (1979) as even attempting to present arguments for such a connection. It’s not that he failed to establish a philosophically interesting indexicality–agency connection—he didn’t even try.

b) Nonetheless, the assumption that such a connection exists is so entrenched that it is worth trying to articulate the arguments we suspect are tacitly (and sometimes embryonically explicit) appealed to.

The chapter is structured as follows. We first present a natural reading of Perry’s paper “The Essential Indexical” according to which it aims to show nothing more than the opacity of action-explanation contexts. Sections 3.2–3.4 make some preliminary remarks about how to articulate stronger (and more interesting) theses about the connection between indexicality (and the de se) and action. We then discuss four possible motivations for these stronger theses. Sections 3.9 and 3.10 outline our positive account of action rationalization and respond to some objections. Our overall conclusion is pessimistic: it turns out that the role of indexicality in explanation and rationalization of action turns out to be minimal and trivial.

3.1. Perry on the Opacity of Explanation Contexts

Despite its familiarity by now, it will be helpful to have in front of us some crucial passages from Perry to draw on. Perry relies on several examples that are structured more or less the same way. One of them is this:

I once followed a trail of sugar on a supermarket floor, pushing my cart down the aisle on one side of a tall counter and back the aisle on the other,
seeking the shopper with the torn sack to tell him he was making a mess. With each trip around the counter, the trail became thicker. But I seemed unable to catch up. Finally it dawned on me. I was the shopper I was trying to catch...I believed at the outset that the shopper with a torn sack was making a mess. And I was right. But I didn’t believe that I was making a mess. That seems to be something I came to believe. And when I came to believe that, I stopped following the trail around the counter, and rearranged the torn sack in my cart. My change in beliefs seems to explain my change in behavior. My aim in this paper is to make a key point about the characterization of this change, and of beliefs in general. (1979, p. 3)

The claim Perry makes about this and the related scenarios is most naturally read as a claim about opacity of explanation contexts. Here are two summarizing passages from Perry where that reading is particularly salient:

At first characterizing the change seems easy. My beliefs changed, didn’t they, in that I came to have a new one, namely, that I am making a mess? But things are not so simple. The reason they are not is the importance of the word “I” in my expression of what I came to believe. When we replace it with other designations of me, we no longer have an explanation of my behavior and so, it seems, no longer an attribution of the same belief. It seems to be an essential indexical. But without such a replacement, all we have to identify the belief is the sentence “I am making a mess”. But that sentence by itself doesn’t seem to identify the crucial belief, for if someone else had said it, they would have expressed a different belief, a false one. (p. 3, bolded emphasis added)

One of the key arguments throughout Perry’s article is this (and variations on it):

Suppose I had said, in the manner of de Gaulle, “I came to believe that John Perry is making a mess”: I would no longer have explained why I stopped and looked in my own cart. To explain that I would have to add, “and I believe that I am John Perry”, bringing in the indexical again. After all, suppose I had really given my explanation in the manner of de Gaulle, and said “I came to believe that de Gaulle is making a mess.” That wouldn’t have explained my stopping at all. But it really would have explained it every bit as much as “I came to believe John Perry is making a mess.” For if I added “and I believe that I am de Gaulle” the explanations would be on par. The only reason “I came to believe John Perry is making a mess” seems to explain
my action is our natural assumption that I did believe I was John Perry and didn’t believe I was de Gaulle. So replacing the indexical “I” with another term designating the same person really does, as claimed, destroy the explanation. (pp. 4–5, bolded emphasis added)

Here is one way to summarize Perry’s claim:

**Indexical Opacity.** There’s a set of indexicals, *I-SET*, that cannot be substituted *salva veritate* in action-explanation contexts by any other expressions.²

It should be clear that this is an instantiation of the more general thesis:

**Generic Opacity.** Co-referential referring expressions cannot be substituted *salva veritate* in action-explanation contexts.

To see that Generic Opacity is true, note that cases analogous to those appealed to by Perry are easily constructed:

**Superman/Clark Kent.** Pushing my cart down the aisle I was looking for CK to tell him he was making a mess. I kept passing by Superman, but couldn’t find CK. Finally, I realized, Superman was CK. I believed at the outset that CK was making a mess. And I was right. But I didn’t believe that Superman was making a mess. That seems to be something that I came to believe. And when I came to believe that, I stopped looking around and I told Superman to clean up after himself. My change in beliefs seems to explain my change in behavior.

We take the S/CK case as showing that in general action explanations don’t have their explanatory force preserved by substitution of co-referential singular terms. Seeing Indexical Opacity as an instance of Generic Opacity suggests that there’s nothing deeply central about indexicals here.

² There are many reasons one might believe Indexical Opacity. For example, you might think that the belief *that I am F*, as had by A, is a different belief from the belief *that A is F*. If an action explanation then cites the first belief, the substitution of “A” for “I” would then change the cited belief, and potentially undermine the explanation.

Our goals, however, are independent of the particular reasons for endorsing Indexical Opacity. We will argue that there is no convincing reason to endorse Indexical Opacity that are not also reasons for endorsing Generic Opacity. So, for example, those who endorse Indexical Opacity because they think A’s belief *that I am F* is a different belief from A’s belief *that A is F* have given no reasons for thinking that this is not just a special case of the general phenomenon that, when A is B, the belief *that A is F* is a different belief from the belief *that B is F*.
A couple of points are worth emphasizing. First, it’s always a bit suspicious to deflate a puzzle by subsuming it under a well-known, larger puzzle, especially when the deflation isn’t accompanied by a suggestion for how to solve the larger puzzle (we are not in this work going to provide a general account of substitutivity failures). Though we recognize that whiff of suspiciousness, we think this particular instance of the strategy is unproblematic. The proponents of Essential Indexicality do not think that they are promoting non-substitutivity as the distinctive feature of indexicals. It’s common ground in this debate that the allegedly distinctive feature of indexicals goes beyond not being substitutable in explanation contexts. That’s why it’s surprising to find Perry summarizing his view in that way. Second, it is of course possible that indexicals exhibit a distinctive kind of opacity (though, note that the passages just quoted provide no evidence of that). Nonetheless, we explore that possibility in Chapter 4 and our conclusion is, again, a negative one.

3.2. Causation, Explanation, Motivation, and Rationalization

Most philosophers (certainly the majority of the more than 1,000 who reference Perry’s classic papers on this topic) take Perry’s arguments to have shown something stronger and more interesting than the opacity of explanation contexts (see, e.g. the quotes we started this chapter with). It is a challenge to articulate precisely these stronger claims and it will help to start with an overview of different kinds of projects in which indexicality can play a role. Consider François. He ducked under the table. There are multiple questions we can ask about that action:

**Causal.** What is the full causal history of François’s action?

**Explanation.** What is the explanation of François’s performing the action?

**Motivation.** What was François’s motivation for performing the actions (what are the reasons for which he performed the action)?

**Rationalization.** What made the action reasonable/rationalizable/comprehensible for François to perform?
The central question, then, is whether any of these questions require answers that cite “de se” or other indexical mental contents. For ease of exposition we propose to focus on Motivation and Rationalization but we think what we say will apply to Causal and Explanation with only slight terminological changes, if you think they should be the focus. 3

3.3. Impersonal vs. Personal Action Explanations/Rationalizations

We are looking for a way to articulate an insight in Perry’s work that goes beyond the opacity of explanation contexts. 4 One way to look for such insights is to look for ways to show that the impersonal

3 We think Motivation and Rationalization present the strongest cases for the proponent of Essential Indexicality, so we focus on those. Plausibly the answer to our central question for Causal and Explanation is “no,” and pretty trivially so. For Causal, we can give a full microphysical history of the world, and we’ll get a causal history of the action that doesn’t cite any de se mental states, because it doesn’t cite any mental states at all. For Explanation, we can observe that explanations are extremely context-sensitive, and that there’s typically no explanatory feature that has to be cited in every explanation of a given phenomenon.

There are potential responses to each of these preliminary points. In response to the Causal worries: If we’re type (or maybe even token) identity theorists, then some of the entities cited in the full microphysical causal history will be mental states (although in a different descriptive guise), and hence may be de se mental states. (Although there are worries about the hyperintensionality of “cite” here.) And if we’re not type/token identity theorists, we may worry that the proposed microphysical history does not in fact give the causal history of the action, but only that of its physical correlates. Here’s a plausible principle: a full causal history of an action must, inter alia, imply that the action is an action. A microphysical causal history absent bridging identities looks like it won’t imply that, though.) In response to the Explanation worries: perhaps we can helpfully appeal to the notion of a full or complete explanation. It’s not obvious that there is any such notion that is sufficiently clear, but arguing that there isn’t looks like more than we want to take on here.

These are all interesting issues to explore in great detail, but doing so goes beyond the scope of this work. In order to bypass these kinds of discussion (that play no role in, e.g. Perry’s or Lewis’s work on this topic), we focus primarily on Motivation and Rationalization.

4 As we have already said, we are not claiming these arguments are found in Perry (1979)—they are rather our attempts to articulate what people have read into that paper. We could have just said: “Perry (1979) does not establish anything but opacity of explanation contexts and so all those who claim he does are wrong and confused,” and leave it at that. But we think that would be too simplistic and uncharitable to those who have read more into Perry. So the goal of the remainder of this chapter is to articulate what we think are the underlying assumptions and arguments that make so many philosophers think there’s a significant indexicality–agency connection (some of the arguments we rely on are inspired by
action rationalizations we are about to cite are incomplete in some significant way. Consider a case to get us started. François thinks François is about to be shot. He doesn’t want François to be shot. So he ducks under the table. Here are two models of action rationalization:

*Personal Action Rationalization 1.*

- **Belief:** François is about to be shot.
- **Belief:** I am François.
- **Belief (Inferred):** I am about to be shot.
- **Desire:** I not be shot.
- **Belief:** If I duck under the table, I will not be shot.
- **Action:** I duck under the table.

*Impersonal Action Rationalization 1.*

- **Belief:** François is about to be shot.
- **Desire:** François not be shot.
- **Belief:** If François ducks under the table, he will not be shot.
- **Action:** François ducks under the table.

For a second illustration of the two models, consider Herman who believes Nora is in danger:

*Personal Action Rationalization 2.*

- **Belief:** Nora is in danger.
- **Desire:** Nora not be hurt.
- **Belief:** If Herman closes the door, Nora will be safe.
- **Belief:** I am Herman.
- **Action:** Herman closes the door.

comments Perry makes in later papers). In other words, if you suspect we are attacking straw-men in what follows, feel free to either ignore the remainder of this chapter (and agree with us that those who think Perry established something of an interesting indexicality–agency connection are all wrong) or come up with better arguments than those we have presented on behalf of our opponents.

5 We are not committed to any particular view of what combination of beliefs, desires, or intentions goes into an adequate action explanation. The points we are making here can be adjusted to whatever your favorite format is for action explanations.
Impersonal Action Rationalization 2.

- **Belief:** Nora is in danger.
- ** Desire:** Nora not be hurt.
- **Belief:** If the door is closed, Nora will be safe.
- **Action:** Herman closes the door.

Note that in this case the impersonal action rationalization doesn’t attribute “Herman”-beliefs or “Herman”-desires to Herman. Instead, the rationalization is entirely third-person. We take our target to be the following view:

**Impersonal Incompleteness Claim (IIC).** Impersonal action rationalizations (IAR) are necessarily incomplete because of a missing indexical component.  

Our guiding question is thus whether there are any arguments to the effect that the IAR-explanations *must* be incomplete. We’ll argue that there are none. First, it is not necessary for an indexical element to enter into the rationalization. Second, on our view the agent doesn’t even need to be represented in a non-indexical way in an adequate action rationalization. Nor does any part of the agent’s body need to be represented (indexically or not). We think there could be a god, who can bring about states of the world just by intending them or maybe just by thinking them. The god thinks, “The door is closed,” and straightaway the door is closed. On our view, this god’s actions can be rationalized even if we don’t specify any kind of *de se* state (or, indeed, even a *de re* representation of the god, or of his body parts).

### 3.4. Preliminaries: Some Weaker Theses and Why We Focus on IIC

We’re inclined to think that the presumed indexicality-agency connection is the source of much of the fascination with indexicality

---

6 Note that in the examples we have focused on an allegedly missing “I,” but the point could generalize: variations will include thoughts or desires involving “here,” “there,” “now,” “we,” etc.
throughout philosophy. It is therefore important to be crystal clear both about what we think the central indexicality–agency connections might be and how the topics of this chapter relate to the remainder of the book.

We will refute IIC. Here’s a reaction we expect to get from many of our readers (and what we got from readers of earlier drafts). You might then be inclined to think something like this: Okay, well, maybe this was all a set-up. Maybe the indexicality–agency connection that matters isn’t IIC. Maybe it’s some weaker thesis. The focus on IIC is just a strawman. First, some remarks about the strawman remark and then some thoughts about some weaker theses that could be advanced. To see that we’re not attacking a strawman, look back at some of the quotes we gave you earlier:

The present suggestion, then, is that indexical concepts are ineliminable because without them agency would be impossible: when I imagine myself divested of indexical thoughts, employing only centreless mental representations, I eo ipso imagine myself deprived of the power to act.7 (McGinn 1983, p. 104, bolded emphasis added)

Kant’s framing his maxim in the first person is no accident. He could not have restated it, for example, as “Immanuel Kant will make lying promises when he is in need.” Such a third-personal thought would not be a maxim of action, since it could not be acted upon by the thinker until he reformulated it reflexively, in the first person. Insofar as the target of universalization is a practical thought, it is essentially first-personal. (Velleman 2006, p. 121, bolded emphasis added)

When I act intentionally on an object, I care neither about the Russellian (Kaplanian) propositions beyond as such nor about their Fregean (Kaplanian) first referents. These are not crucial. What I need is to bring those doxastic referents somehow into my experience through my thinking indexical (that is,

7 Here is the sentence immediately preceding this passage: “As I remarked in chapter 5, indexicals have recently been associated with practical reasoning, the thesis being that indexical thought is a necessary condition of agency” (p. 104). And here’s what McGinn says in ch. 5, “And as a number of writers have observed, dispositions to action induced by indexical thoughts correspond with the linguistic meaning of the associated indexical sentence: you and I are disposed to act in the same way if we both think ‘I am about to be attacked by a bear’ (other things being equal)” (p. 65). Both quotes are immediately followed by a citation of Perry 1979.
experiential) references...the singular proposition that mobilizes my powers of deliberation and action is the indexical truth within my grasp. (Castañeda 1989, p. 126, bolded emphasis added)

It is widely agreed that agents need information in an egocentric form: they must think of places as "here" and "there", times as "now" and "then" if they are to be able to act on what they know (Perry 1979). (Owens 2011, p. 267, bolded emphasis added)

We take IIC to be a fair summary of these kinds of claims (and we could have added many, many more citations⁸). Now you might think these claims overreach. Maybe they should have been more modest and so charity requires that we look also for some weaker indexicality–agency connections. What might such weaker connections be? Consider first the following claim:

**Impersonal Incompleteness Claim 1 (IIC1).** In some action explanations/rationalizations, indexicals occur ineliminably.

Important point: We don’t know how to read IIC1 as anything but a trivial corollary of the opacity of action explanations/rationalizations. In other words, we grant it, but it shows nothing distinctive about indexicals. For any content, C, there are contexts in which substituting C for any other content can result in a change in truth value. There are, for example, action-explanation (rationalization) contexts in which you can’t substitute “Superman” for some other co-referential term. In the same way there are contexts in which substituting an occurrence of “I” for a co-referential term can’t be done salva veritate. None of this should be a surprise and we don’t need to read papers by Perry and Lewis to learn about it. Recall, we assume throughout that (i)–(ii) are in the common ground among participants in this debate:

1) Belief attribution and action explanations exhibit opacity.
2) Indexicals have non-stable characters.

IIC1, as we understand it, commits one to nothing more than that. Now, you might think that non-substitutivity as it applies to indexicals is distinctive in some way. If you think that, it is very important to keep

---

⁸ See, for example, the passages from Perry quoted in Sect. 3.6.
the structure of these first chapters in mind. In Chapter 4 we explore and then refute the thought that the behavior of indexicals in opaque contexts is distinctive in some way. So if that’s your reason for interest in IIC₁, hold off until Chapter 4 (or maybe go directly there). In sum, we conclude that IIC₁ is true but uninteresting;⁹ to see the full argument for this conclusion, though, you’ll have to look at this chapter in conjunction with the next.

If you agree with that response to IIC₁, you might be tempted to look for a different weaker version of IIC—one that isn’t just about opacity but also isn’t as strong as IIC. We can imagine various versions of IIC₂:

**Impersonal Incompleteness Claim 2 (IIC₂).** As a matter of deep physical necessity, indexical beliefs are needed in order for human beings, constituted as they are, to act.

We talk more about this reply at the end of this chapter (see Objection 4), but in brief: none of the arguments (or motivations) we are going to consider have IIC₂ as a potential conclusion. That should not be surprising—none of the arguments have facts about the physical structure of humans as *premises*. There is no attempt in the arguments to study in detail the underlying physical structure of humans and their ability to act. That would require arguments and evidence of a completely different kind from what we find in the philosophical tradition we engage with in this work.

So our strategy is to engage with IIC. We should note right away that this isn’t an easy task since there are few real arguments out there to engage with. Put simply, the remarks prominent philosophers make about the indexicality–agency connection don’t seem very plausible candidates for establishing IIC. As a result, we structure the remainder of the chapter around what we call “Motivations.” They’re not really arguments, but they are observations that seem to us to play an important role in the widespread endorsement of IIC. There is a danger here that we are very much aware of: suppose A endorses a thesis *T*, but doesn’t tell us why. We then try to figure out why she endorses *T*. If in

⁹ Again, on the assumption, which we are granting *arguendo*, that there are opaque contexts.
doing so, the candidates we come up with all fail to establish $T$, there’s the immediate suspicion that we’ve been uncharitable and that either there are some other (and better) reasons for endorsing $T$ or that $A$ didn’t endorse $T$. The second option we just dealt with. With respect to the first option: we would be delighted if someone could respond to this book by constructing better arguments for IIC. That would really move the debate forward.

3.5. Motivation 1: Judgments about Cases

The cases at the beginning of Perry’s “The Essential Indexical” are intriguing, ingenious, suggestive, and have for good reasons captured the imagination of at least two generations of philosophers. We suspect that reflection on these cases motivates many of those endorsing various indexicality–agency connections. But what can we learn from them? One thing to note right away is that these cases are not like what has become known as “the Gettier case” or Burge’s famous “arthritis case” in his paper, “Individualism and the Mental.” Burge and Gettier presented counterexamples to necessity theses (put simplistically, the claims that necessarily all instances of justified true belief amount to knowledge and necessarily all contents are determined by “what’s in the mind of the speaker or thinker”). Burge’s and Gettier’s cases, if they succeed, show that these necessity claims are not true. This is not the dialectical setting of Perry’s cases. There are two opposing views: on one view, a distinctly first-person (or other indexical) attitude is required in an adequate action rationalization. According to the other view, it is not required. Perry’s cases have the form of thought experiments in which it is stipulated that at $t_1$ there is no action (and no relevant first-person or other indexical attitude) and then at $t_2$ there’s action. By stipulation, the only difference between $t_1$ and $t_2$ is the addition of a first-person state. So we all say: fine, yes, that could happen—the first-person state could have that effect and so could play an important role in rationalization of that particular action. Note that this is at best super-weak inductive support for a strong thesis such as IIC. Without further argument, this won’t justify a necessity claim like IIC. That an ignorance of the kind the agent starts out with in Perry’s examples can block
rationalization of agency, doesn’t show that it has to. Cases can be used as counterexamples to universal claims, but can’t be used to (deductively) establish the latter.  

3.6. Motivation 2: Bodily Movements Require Indexical Thought

Here is a passage from Perry and Israel’s “Fodor and Psychological Explanation”:

…the need for context-sensitive representation never goes away. If he wants to drink the coffee, Jerry still has to pick up the cup in front of him.

Consider, for a different example, Jerry’s making a phone call to Zenon. Here the belief, say, “Zenon’s phone number is 555-5555” is relatively context-insensitive. There is a present-tense marker, but phone numbers are relatively stable properties of persons, so we can ignore that. When Jerry calls Zenon, the perceptions that originally gave rise to the belief may be remote in time, and the effect of the action that fulfills the goal—the ringing of Zenon’s telephone and his answering it—are remote in space. Neither the number nor Zenon needs to be thought of demonstratively or indexically to understand the transaction. But, like all human action—except perhaps pure ratiocination—the crux of the matter involves physical interactions with a physical object. The practical reasoning will terminate with some context-sensitive way of thinking about the telephone. The Fodorian version of half of Kant’s maxim about concepts and intuitions should be “Eternal tokens of mentalese without context-sensitive tokens of mentalese are blind” (1993, p. 311, bolded emphasis added)

In a later paper (1998), Perry says:

Consider a transaction with a fax machine. To press certain buttons on it, I have to move my fingers a certain distance and direction from me. It isn’t enough to know where the buttons were relative to one another, or where the
fax machine was in the building or room. I had to know where these things were relative to me (p. 87, bolded emphasis added)\(^1\)

Note right away that this isn’t really an argument with premises and a conclusion. It is more of a powerful and persuasive statement of the view we’re opposing. That said, we suspect this kind of rhetoric can be effective. We suspect one might have this reaction: Right, isn’t that just obvious?

Reply to Motivation 2. Here is why you shouldn’t be moved by the kind of rhetoric we find in the Perry quote: in order to act, our bodies have to be brought into engagement with the world around us. That much is trivial. But what is not trivial is that for that to happen, we must represent ourselves in relation to the objects we engage with. Consider what Perry says about the fax machine: why do you need to know or have a belief about where the buttons are relative to you? In general, we don’t need to have knowledge or beliefs of all the facts involved in our action. To move a finger, you have to tense various sequences of muscles. But you don’t need to know or have beliefs about what those sequences of muscles are. Somewhere in the physical architecture, the relation of you to the muscles gets implemented, but it doesn’t need to be a cognitive implementation. There is no reason why it can’t be like that for you and the buttons. On the cognitive–representation level, it’s just all about the buttons and their objective position in space. Then a bunch of neurons fire and our bodies end up doing the right thing. If there is no representation of the fingers (whether conscious or subpersonal), there is a fortiori no crucial indexical representation of the fingers. Everyone agrees that the representational level gives out somewhere. For Motivation 2 to be persuasive, it needs to come with some reason to think that it gives out at exactly the right spot: while the self is still involved in the representation. But the passage from Perry doesn’t even make an effort to develop an argument to the effect that this it is impossible for representation to end elsewhere—no evidence is presented that the action couldn’t take place the way we just sketched.

\(^1\) For those inclined to doubt that IIC is endorsed by our opponents (despite the passages we have quoted), note especially the use of “had to,” “have to,” and “the need for context sensitive representation never goes away” in these passages.
A couple of further points to bolster the non-indexical characterization of the action: first, note that Perry’s claim about the need for a representation of the agent is, on a simple reading, obviously false. It is not true that the agent and the buttons have to be brought into engagement. What is true is that the agent’s finger and the buttons need to be brought into engagement. That, however, is just two objects in the world that need to be coordinated. There is no good reason to think that she must know where she is relative to these events. Typically, people are near where their fingers are, but in weird cases, that could change, and it wouldn’t matter to the nature of the action.

Second, note that if she happens to think about the buttons’ location relative to her, no argument has been given that she has to think about herself in a distinctly first-person way (whatever that means). If the agent is Nora and she happens to represent Nora’s relationship to the buttons, she could represent herself in a non-first-person way, e.g. as Nora.

In sum, there are at least two mistakes underlying Motivation 2.

**Over-Representation Fallacy.** Because the body is involved in the movement, the body needs to be mentally represented.

**Relational Fallacy.** If some part of the body has to be represented, it has to be indexically represented.

There simply are no arguments or evidence for either of these claims. Not everything involved in an action needs to be represented by the actor. And even if you maintain that some part of the body has to be represented, it doesn’t follow that it has to be indexically represented. What’s important is just that it gets represented, so that it can be directed or controlled. Indexicality just seems to be irrelevant.

It is helpful to contrast our points here with one made by Millikan. Millikan—one of our few allies against Essential Indexicality—says:

Now it is trivial that if I am to react in a special and different way to the knowledge that I, RM, am positioned so in the world, a way quite unlike how I would react knowing anyone else was positioned so in the world, then my inner term for RM must bear a very special and unique relation to my dispositions to act. But what does that have to do with indexicality? My inner name “RM” obviously is not like other names in my mental vocabulary. It is a name that hooks up with my knowhows, with my abilities and dispositions to act, in a rather special way.... My inner “RM” is indeed special. Let us call
it @”RM”, or RM’s “active self name”. It names a person whom I know, under that name, how to manipulate directly; I know how to effect her behavior. But in order to know how to manipulate this person, why would I need to think indexical thoughts? What has knowhow to do with indexicality? (1990, p. 730)

Millikan is right, and importantly so, in pointing out that there’s not even the beginning of an argument for anything properly labeled “essential indexicality” here. If anything is shown, it has to do with something like what Millikan calls an “active self-name” and it need not be indexical. However, we think that even the claim about an essential role for the “active self name” (or “de se concept,” if you prefer that label) is unsupported. Why think that the agent has to represent herself at all in order to push the button? What we get is simply a claim that this is required, not evidence or argument to the effect that the act couldn’t happen without a representation of the agent.

3.7. Motivation 3: Generalizations about Actions

At the end of “The Essential Indexical,” Perry presents a positive view about indexical content. Perry summarizes his view as follows:

I propose we look at things in this way. The shoppers, for example, are all in a certain belief state, a state which, given normal desires and other belief states they can be expected to be in, will lead each of them to examine his cart. But, although they are all in the same belief state (not the same total belief state, of course), they do not all have the same belief (believe the same thing, have the relation of belief to the same object). We use sentences with indexicals or relativized propositions to individuate belief states, for the purposes of classifying believers in ways useful for explanation and prediction. That is, belief states individuated in this way enter into our common sense theory about human behavior and more sophisticated theories emerging from it. (1979, p. 18)

The following argument can be extracted from this passage. Explaining François’s action using “François”-beliefs rather than “I”-beliefs fails to capture important generalizations. We don’t explain why François and Josh both act in similar ways in similar circumstances. When someone
has a belief at \( t \) that would be expressed by the believer (were he an English speaker) by an utterance at \( t \) of the sentence type “I am about to be shot,” that person ducks under the table. A bit more generally: suppose the Impersonal Action Rationalization model can succeed. Now suppose that there is some rationalization \( R \) of François’s action, featuring various “François”-beliefs and -desires, but no indexical beliefs and desires. Here’s a claim: any agent \( A \) who had the analogous set of beliefs and desires (that is, beliefs and desires with “François” uniformly replaced with “\( A \)” ) would be rational to act in the same way as François. The Impersonal Action Explanation model doesn’t explain that. The Personal Action Explanation model does explain it: on this model, exactly the same explanation (involving indexical beliefs and desires) applies to both agents.\(^{12}\)

Reply to Motivation 3. Note that the argument we just sketched didn’t end with a conclusion. One conclusion one might draw is this: There are some generalizations that impersonal action explanations cannot capture. Note that this is very far from an interesting thesis like IIC. Suppose it is true that there are some generalizations involving indexicals that help us classify “believers in ways useful for explanation and prediction” (in Perry’s phrase). That’s compatible with there being other, non-\( de \ se \)/first-person generalizations that also are useful in that way. There are all kinds of generalizations we might want to capture, some metalinguistic, some not (and this bolsters our claim that generalizations that appeal to allegedly special \( de \ se \) content have no special status). Here are two generalizations that do not involve indexicals and that, at least for some purposes, will help us classify “believers in ways useful for explanation and prediction”:

**Useful (U1).** Any criminal engaged in a criminal act who has a belief she would express by “Superman is right around the corner” will, \( ceteris paribus \), run.

\(^{12}\) Not every story about indexical contents or \( de \ se \) states will yield the result that the Personal Action Explanation captures the generalization by attributing the same mental states to François and to \( A \). The crucial question is whether François’s first-person belief that he is \( F \) has the same content as \( A \)’s first-person belief that he is \( F \). Accounts vary in their answer—Fregean pictures of first-person thought, for example, tend to attribute different contents to François’s and \( A \)’s first-person thoughts, while Lewisian accounts attribute the same content.
Useful 2 (U2). For any person, \( x \), if there’s a button, \( y \), that \( x \) wants to push, \( y \) is close to \( x \), \( y \) is within \( x \)’s visual field, and \( y \) is easily within \( x \)’s reach, then \textit{ceteris paribus}, \( x \) will push \( y \).

Given the immense contextual flexibility in what counts as useful for explanation and prediction of action, there are very natural settings in which \( U_1 \) and \( U_2 \) can serve that purpose at least as well as a \textit{de se}-involving generalization such as \( U_3 \):

Useful 3 (U3). Anyone who has a belief she would express using “I am being pursued by a bear” and believes she can avoid being eaten by that bear by climbing up a nearby tree, will, \textit{ceteris paribus}, climb up the nearby tree.

To re-emphasize the main point here: even if we grant that generalizations like \( U_3 \) are true and in some significant sense involve indexicality, that’s too weak to support IIC.  

3.8. Motivation 4: Anti-Generalizations about Actions

The previous argument focused on the need to capture sameness of beliefs and desires among agents who are motivated to act in the same way. Now we focus on the need to capture difference in belief and desire among agents who are motivated to act differently. Consider a case where two agents have all the same objective, or non-indexical, information about the world. Not only can they be motivated differently, but what it is rational for them to do can differ. To see what

\[13\] This is a point at which several of our readers have found that they are unsatisfied with our focus on IIC. They agree that Motivation 3 fails to establish or support IIC, but find that it is still interesting that there are generalizations about action involving indexicality. We attempt to address this concern in Sect. 3.4. To summarize: if you think that these generalizations are not just a manifestation of Generic Opacity, and the fact that any different expressions will give rise to different generalizations, but are evidence for a form of Essential Indexicality other than IIC, then you need to formulate the appropriate Essential Indexicality thesis. We are unable to see that there is an interesting Essential Indexicality thesis supported by Motivation 3.
the proponent of IIC might have in mind, consider this example from Perry:

I shall use the term “locating beliefs” to refer to one’s beliefs about where one is, when it is, and who one is. Such beliefs seem essentially indexical. Imagine two lost campers who trust the same guidebook but disagree about where they are. If we were to try to characterize the beliefs of these campers without the use of indexicals, it would seem impossible to bring out this disagreement. If, for example, we characterized their beliefs by the set of “eternal sentences,” drawn from the guidebook they would mark “true”, there is no reason to suppose that the sets would differ. They could mark all of the same sentences “true”, and still disagree in their locating beliefs. It seems that there has to be some indexical element in the characterization of their beliefs to bring out this disagreement. (1979, p. 5)

Perry’s focus here is on how to bring out the disagreement between the two agents. We cannot do so without appealing to indexical content, he claims. We want instead to focus on a separate point: what it is rational for the two agents to do will differ. They both want to reach Mount M. For one of them, it will be rational to turn left, for the other to continue straight ahead, even though they have the very same non-indexical beliefs.

Reply to Motivation 4. As it is stated, this argument doesn’t conclude with anything as strong as IIC. Even if it were true that in some case what makes it rational for A to perform an action, but not for B, is that A has an indexical belief that B does not have, nothing general follows about the connection between indexicality and agency. All kinds of difference in beliefs can make a difference to what it is rational to do. The fact that action explanation and rationalization contexts are opaque makes this particularly obvious. Consider the following case:

A and B are being threatened by Lex Luthor. They both see a person in the distance. A believes that the person is Superman; B believes that the person is Clark Kent. Now A has reason to refuse to give in to Lex Luthor’s threats, since he has reason to think Superman will save him. But B has reason to submit to Lex Luthor’s threats, since he has reason to think Clark Kent won’t save him. A and B have all the same objective information and desires here, but have reason to act differently.

Notice that there’s nothing distinctively de se and nothing indexical going on in this case. So the Perry phenomenon can’t in itself be an
argument for the distinctiveness of indexicals. At most it can establish
that there is a way that A is thinking of A that isn’t the way that B is
thinking of A. But that’s no more than a familiar instance of opacity,
nothing distinctive about indexicals.

It might be tempting to try to improve on this argument for IIC in
the following way. In any case where someone, A, is characterized only
by her non-indexical beliefs, we can imagine another agent, B, with the
same non-indexical beliefs, but rationally motivated to pursue a different
course of action because of the presence of an indexical belief not shared
with A. In reply to this attempted improvement, we say: even if this were
ture, it doesn’t show that as a matter of fact indexical beliefs are always a
component of what rationalizes an action. If this isn’t obvious, note that
the following may be true (or in any case, we have as much reason to
believe it as we do the corresponding claim about indexical belief): in
any case where someone, A, is characterized only by her non-Superman
beliefs, we can imagine another agent, B, with the same non-Superman
beliefs, but rationally motivated to pursue a different course of action
because of the presence of a Superman-involving belief not shared with
A. For example, suppose A and B are alike with respect to their beliefs
and desires about what to order for dessert: it all points towards apple
crumble as the rational choice. Now imagine that B has the added belief
that Superman will torture everyone who eats apple crumble for des-
sert. The addition of this belief will make a difference as to what it is
rational for each to do: it is rational for B to not order the crumble, while
it is rational for A to order the crumble. This point obviously generalizes,
but that’s not a good reason for thinking that “Superman”-beliefs play dis-

3.9. The Selection Problem and a Positive Proposal: The Action Inventory Model

So far this discussion has pointed out flaws in arguments attempting to estab-
lih IIC. Someone convinced by those criticisms might still wonder why it
is the case that “when you and I both apprehend the thought that I am
about to be attacked by a bear, we behave differently” (Perry 1977, p. 494).
More generally, an agent will have indefinitely many third-person beliefs
Indexicality, the \textit{de se}, and Agency

and desires about the world that don’t result in or rationalize action by the agent. So we need to explain which ones produce action and which ones don’t; we need a filtering mechanism of some sort. Call this “the Selection Problem.” The \textit{de se} proposal is that the first-person status is the filtering mechanism; only the first-person beliefs and desires give rise to the right sort of intentions, which then gives rise to action/rationalization of action. What can be said about the Selection Problem if we don’t appeal to \textit{de se} states? Here’s a toy picture that we offer as a replacement for the \textit{de se} picture:

**The Action Inventory model.** Every agent has a very wide range of third-person beliefs and desires that give rise to third-person intentions, which in turn rationalize or motivate actions (via their recognition). Not all of these intentions are going to produce action, at least in normal cases (perhaps in a god they would). This is because a given agent has an “action inventory”: a range of actions that he can perform. An agent constantly seeks to match his intentions with his action inventory, and when he finds a match, action occurs. When there’s no match, the intention idles, and doesn’t motivate or rationalize action. So the Selection Problem is solved by appealing to the physical or psychological constraints of the agent: only certain actions result because only certain actions were available in the first place.\textsuperscript{14}

Now, this is going to look like a cheat if we think that the actions have to be specified first-personally—such as “I move my hand thus-and-such” or (without specification of subject, as in Lewis) “kick my left leg.” Then the thought will be that there’s first-person content in the action inventory, and it’s only via the appeal to those contents that we solve the Selection Problem. But there’s no reason we need to think of it that way. The action inventory can include “John throws a baseball through a window,” “That hand lifts a glass,” “The car steers out of the skid,” or even “Kripke releases his unpublished papers” (in the inventory of someone other than Kripke). The thought is that we’re embedded agents who have a range of capacities directly to manipulate things in the environment. Some of those things are ourselves and our parts, but others are

\textsuperscript{14} We hope it is obvious that we are taking the Action Inventory model pretty loosely—we’re not particularly committed to anything like a serious model with a switchboard in the head. We don’t need that this model is really right at all, but just that it’s a possibility.
frequently and infrequently encountered objects, both proximal and distal. To think otherwise is to pursue the “how to move my body” line of thought, which we have already argued against. According to our alternative picture, the belief-desire-obligation-intention sets produce a bunch of inputs—potential actions that are ready to go. Those inputs (i.e. potential actions) then hit the “action center,” which is a big switchboard with a bunch of available actions. If an input matches an available action on the switchboard, an action results. But no one has to “look” to see if there’s a match, and hence no one has to “think about the available actions” in any way, let alone a first-person way. The actions have already been “thought about” by the time they emerge as potential actions (and thought about just as the third-person action “that \( p \)’); all that remains is to see if the actions are among the things that can be done.

One way to put our basic point is to say that the Selection Problem is solved by identifying a category of actionable contents. Most likely there isn’t much systematic and general to say about that category. What constitutes actionable content for an agent will depend on the kinds of sensory inputs the agent has—for example, whether he moves around the world only with smell and taste or using sonar and sounds, eyes and ears and fingers, etc. It will depend on the agent’s physical setup: whether he has fingers, claws, a trunk, wings, feet, fins, etc. It will depend on whether the agent’s skills include steering cars out of skids or balancing plates on poles, and whether the agent’s casual reach includes having employees publish Kripke’s manuscripts. It is unlikely that there is much to say in general about all content that combines with the propositional attitudes, intentions, and other cognitive states, perceptual inputs, etc. of such varied agents to trigger action. Despite this variability, two negative points are worth emphasizing:

1) It is not necessary that the agent believes or knows that they are within actionable reach. There has to be some part of the agent’s action-generating architecture that recognizes the contents as representing potential actions that are within actionable reach, but that doesn’t have to rise to the level of a belief or be represented as being within actionable reach.

2) Sometimes the actionable contents are represented as being within actionable reach, the agent believes that they are, but in
those cases they are representations of the relationship between the agent and the actionable content. The agent doesn’t have to think of him- or herself, the actionable content or the relation between him- or herself and the content in any special way. They are simply thoughts of the form, “agent $a$ is actionably related to object $o$.”

3.10. Clarifications, Objections, and Replies

Objection 1. What you say here is incompatible with basic elements of folk-psychological explanations of actions. According to your view, two people can have the same beliefs and desires (the belief that François is about to be shot, the desire that François not be shot, etc.), but act and be rationally motivated to act in different ways. This is an unacceptable consequence.

Reply. This objection mischaracterizes our view (which is easy to do since it requires thinking in a somewhat unusual way). On our view, the actions will differ, but not what we are rationally motivated to do. Someone other than François, e.g. Dilip, can be rationally motivated to perform the action that François climb into the tree. Dilip just can’t perform the action. François can perform it, so he does. François is, like Dilip, rationally motivated to perform the action that Dilip call for help. But François can’t perform that act. So François climbs the tree and Dilip calls for help. Now, if you think that folk psychology requires that when two agents have the same belief-desire sets they perform the same action, then we are denying that. But surely that’s not a view worth preserving. A more plausible principle is that ceteris paribus, they act the same. That principle we do endorse. It is because of the difference in their available actions that François and Dilip don’t act the same. Everyone should agree that you can’t explain what someone actually does just by talking about their psychological states. They can have all the psychological states you want, but if they don’t have hands and legs, for example, there are lots of actions they aren’t going to perform. Presumably everyone wants these

---

15 Dilip Ninan has forcefully pressed this concern on us.
sorts of cases to be swept into the *ceteris paribus* clauses. We want a bit more than that. There’s a legitimate debate to be had about how broad the *ceteris paribus* clauses should be in the ideal theory of action, but we see no reason to think that folk psychology as such speaks on where to draw the boundaries.  

**Objection 2.** What about present tense? Maybe agency doesn’t require first-person beliefs, but surely “now”-thoughts are required. Otherwise, it would be a mystery how action could take place. Consider an agent about to turn on the light by flipping a switch. She wants the light to be on, she knows that flipping the switch will turn on the light, she intends to flip the switch, and believes that moving her hand and then finger in a certain way can make it happen. But, the objection goes, these are not tenseless intentions, beliefs and desires. She wants the switch flipped soon (close to now), which is why she now starts moving her hand, and now moves her finger. If she didn’t think, Now the finger is within the reach of the switch, why would she move it then, and not at another time?  

16 A comparison with Fregeanism vs. direct reference theory might help here. Fregeanism claims to be better off than direct reference theory because it can explain, e.g. inferences in terms of contents. The difference in sense between “Superman” and “Clark Kent” explains why inferences from “Superman is *F*” and “Clark Kent is *G*” to “Superman is *F* and *G*” aren’t made. Direct reference can’t give such explanations. This looks like a win for Fregeanism, but then we note that Fregeanism can’t give us error-proof explanations: minimally, there’s a “rational agents” proviso; furthermore, there’s just no plausible picture of senses in which there’s no possibility of misidentification of sameness of sense. So both direct reference theorists and Fregeans will have to say at some point that there are just brute facts about whether we “hook things up” mentally. So if you’re looking for error-proof explanations, you’re not going to get it, and if that was your reason for wanting Fregeanism, you should give up on it. Similarly here: no one’s going to get infallible explanations of action, in which having the right mental states just guarantees that everything works out. We’re all going to have explanations which sometimes result in error. See Fodor (1995) for a defense of this line of thought.

17 If you’re a functionalist about mental content, you might think that you just can’t have the sort of story we have, since on that story mental contents are (roughly) just individuated by their potential to produce action. So if François and Dilip act differently, that already establishes that they believe differently. We will return to a detailed discussion of functionalism in Ch. 6, but here we just note that the functionalist point can’t be made quite so easily. Any plausible functionalist story will also need to include various *ceteris paribus* clauses, or its picture of content will bear no recognizable relation to our pre-theoretic mental state ascriptions. So what’s needed here is again an argument that the degree of error-tolerance in the functionalist story needs to be exactly what our opponents say it is, rather than what we suggest it could be.
Reply. Recall the Action Inventory model: our belief-desire-obligation-intention sets produce potential actions that are ready to go; they encounter a set of available actions. If an input’s content matches the content of an available action, an action results. Since no one has to “look” to see if there’s a match, no one has to “think about the available actions” in any way, let alone a de nunc way. Pre-theoretically, this is unsurprising: the actions we care about are complex, consisting of an indefinite number of sub-actions; moving one’s hand, stretching one’s finger, touching a switch, pushing down on one’s finger, etc. To think that each of these will have to be represented not just as an initiated sequence, but that each moment of transition into a new stage has to be represented by a belief-desire-obligation-intention set involving “now” is to attribute unnecessary and baroque representational powers to us.

Objection 3: “Consider this dialogue:

A: Why did Francois duck?
B: Because he didn’t want Francois to be shot, and he thought that if Francois ducked under the table, Francois would not be shot.
A: Wait—did Francois realize that he himself was Francois?
B: I’m not sure. And it doesn’t matter anyway: I already explained to you why Francois ducked.
A: No, you didn’t. For suppose Francois had believed that he was not Francois, but that the man standing beside him was Francois. Then surely, given the beliefs and desires you said Francois had, he would have done something like shout “Duck!” at the man beside him rather than duck himself. What Francois would do given the beliefs and desires you said he had depends on who he thinks Francois is, and how he believes that he is related to the person that he thinks is Francois.

On your view, it seems like B’s second response is adequate, and that there is something wrong with A’s last remark. But what A says seems entirely reasonable. What do they think is wrong with A’s remark?”

18 This objection is from an anonymous reader for OUP.
Reply. Maybe the easiest way to see why this response is misguided is to recall that we started out with two Impersonal Action Explanations. One involved François and his “François”-beliefs and -desires. The second concerned Herman and his “Nora”-beliefs and “Nora”-desires. To see where this objection goes wrong, focus on the second case. In that case, Herman wasn’t motivated to act for reasons having to do with Herman and beliefs about Herman didn’t figure into the explanation or rationalization. It was all about Nora. In such cases, the agent’s possible identify confusions are irrelevant.

Objection 4: Suppose I concede that action could proceed non-indexically. I am really doing a form of high-level empirical work on human action. My conclusion is that it’s a contingent (but maybe physically deep) feature of humans that we act only indexically, because we’re embodied and have to move our bodies. I am doing a form of psychology or cognitive science about very general truths about humans (which is why I don’t have to do actual empirical work). In short, your refutation of the necessity claim leaves open that there’s a deep and significant contingent connection between indexicality and rational agency.

Reply. One way to see why this reply is hopeless is to focus on the parenthetical remarks. There’s no such thing as “doing psychology about very general truths about human beings” without empirical backing. The kind of non-empirical approach found in the Perry–Lewis tradition we are engaging with cannot establish contingent truths about how humans act, how they represent the world around them when they act, how they move their bodies, or how, as a matter of fact, they rationalize their actions, no matter how general. Whether indexicality (or “the de se”) is involved is a very, very detailed question about the implementation of complicated mechanisms in the human head. Armchair reflections about us moving our fingers won’t get us such conclusions; nor will philosophical reflections about generalizations or opacity. In saying this we are not taking a stand on how action mechanisms are in fact implemented in human heads. It would be absurd for us to think that such a complicated empirical issue can be settled that way. We’re just saying that our opponents don’t know how
the implementation goes, and that there are feasible non-indexical implementations.\textsuperscript{19}

\textit{Objection 5. This is all very unfair to Perry (1979). He didn’t try to argue for IIC. Perry’s goal was to refute what he called “the doctrine of propositions” (p. 6). Even if your arguments against IIC succeed, it doesn’t touch Perry—he still succeeds in his stated aim.}

\textit{Reply. We should remind the reader of comments made earlier. It is widely assumed that Perry didn’t just present arguments in favor of the opacity of explanation contexts or against one interpretation of Frege’s view of propositions. The philosophical tradition has taken Perry’s papers (and his 1979 paper in particular) as establishing that some important connection exists between indexicality and agency. This chapter has attempted, on behalf of that tradition, to spell out what that connection might be and how it might be argued for. We agree that our pessimistic conclusions don’t show that Perry failed in his stated goals—maybe the charitable reading of Perry is that he had no interest whatsoever in establishing an interesting indexicality–agency connection. If so, the tradition has been mind-blowingly careless in their Perry exegesis.}\textsuperscript{20}

\textit{Objection 6: You shouldn’t be focused on IIC—it’s a red herring.}

Several readers of earlier drafts ended up agreeing with us that IIC is false and that no good arguments or motivations can be found for it. Some reacted by asking: \textit{Why focus just on IIC?} We end this chapter by reminding readers with that reaction to look back to Section 3.4 where we justify our focus on IIC.

\textsuperscript{19} This is true whether or not you are a functionalist about content. In both cases, stories about contingent features of the theory of content are beholden to a wide range of empirical facts that are never examined in the arguments.

\textsuperscript{20} It is, of course, not entirely unnatural to read Perry as trying to establish some indexicality–agency connection. All of Perry’s examples concern the transition from not acting to acting, and there is a great deal of discussion of action explanation in the paper, which is not just about action-explanation terminology or contexts.
3.11. Brief Summary of This Chapter

There are six central take-home messages from this chapter:

1) The thought that first-person (or indexical) thought has a philosophically significant connection to agency is, without any doubt, a powerful one. It is endorsed by a very wide range of excellent philosophers working across all fields of philosophy. It is interesting and surprising, therefore, that it is difficult to find any precise articulation of that connection and even harder to find arguments for it.

2) On at least one natural reading of the arguments in Perry 1979, they establish nothing more than that indexicals can’t be substituted \textit{salva veritate} in explanation contexts. The claim that explanation contexts are opaque, however, is not surprising. Those drawn to the thought that agency and the first person are significantly connected want more than that.

3) What they say they want is some version of IIC (and it is hard to see any weaker thesis that would be of interest).

4) What arguments can be given for IIC? Here we are left with a bit of a challenge since there are hardly any genuine arguments to be found in written literature (though, in conversations and Q&A’s, we have encountered quite a few). In light of the lack of arguments, we engaged in a bit of philosophical reconstruction: we tried, on behalf of our opponents, to articulate what we suspect are the (often tacit) motivations for endorsements of IIC.

5) We show that none of those motivations come close to establishing IIC.

6) Finally, we sketched an account of action explanation/motivation that did not give any kind of privileged role to indexicality or first-person thought.
Chapter 4

Indexicality, Opacity, and Fregeanism

In Chapter 3 and in later chapters we repeatedly make the following kind of move: in response to some allegedly distinctive feature of indexicals, we say that it is just an instance of opacity. We shouldn’t be surprised or think it distinctive of indexicals that they can’t be substituted salva veritate with co-referential expressions in, e.g. belief contexts. If a context is such that substituting co-referential terms isn’t guaranteed to preserve truth value, then substituting an indexical referring to $o$ for a non-indexical referring to the same object $o$ can result in a change of truth value. This shows nothing distinctive about indexicals and provides no evidence for any kind of Essential Indexicality thesis.

Maybe this is too quick. Maybe there are different kinds of opacity and maybe the opacity exhibited by indexicals is distinctive in some way. Maybe the kind of failure of substitutivity we face when indexicals are involved is not just of the kind that we find in off-the-shelf cases involving “Superman” and “Clark Kent” or “Hesperus” and “Phosphorus.” If this were the case, considerations involving indexicals in opaque contexts could provide arguments for a revised notion of content—it could motivate a theory of content in which indexical contents are distinct in ways other than simply having variable characters (in Kaplan’s sense).

To explore these topics, we focus on two claims:

**Motivational Distinctiveness.** There are reasons for thinking that indexicals interact with opaque contexts in ways that are unique to indexicals, which can’t be replicated for non-indexical expressions.
Diagnostic Distinctiveness. There are reasons for thinking that our account of opacity, e.g. names can’t be extended to indexicals.

Since the question of whether Diagnostic Distinctiveness is true is potentially enormously broad (requiring us to look at every proposed account of opacity), we limit our discussion to how contemporary followers of the German philosopher and mathematician Gottlob Frege (so-called Fregeans) deal with indexicals. So, the version of Diagnostic Distinctiveness we will primarily be concerned with is this:

Fregean Diagnostic Distinctiveness. Within a Fregean framework, the explanation/theoretical account of opacity for names can’t be extended to indexicals.

Our conclusions are these:

• The kinds of cases that have been made famous by Perry and others pattern with (indeed, are instances of) standard Frege cases and don’t even provide prima facie support for Motivational Distinctiveness.
• We then consider Fregean Diagnostic Distinctiveness and show that when we look at the details, it turns out that, e.g. Perry’s reasons for thinking that the Fregean story can’t be extended to indexicals are just Kripke’s reasons for thinking that Fregean stories of any sort won’t work. So that’s not a special reason to think that indexicals cause problems.

In sum: when we look at the interaction between indexicals and opaque contexts, there’s just the general problem of Frege puzzles. That’s a hard problem, and maybe no one has a good solution to it, but no evidence has been provided that we need to be revisionary about content to accommodate indexicality.

Before launching into the case for our conclusions, a word on how they fit into the broader dialectic: in other chapters, one (although not the only) move we make against proposed forms of Essential Indexicality is to observe that the putative essentiality seems to flow directly from certain kinds of opacity. So if we are successful in this chapter in making the case that Motivational and Diagnostic Distinctiveness are false, or at least unsupported, then that move elsewhere goes through.
But even if we are unsuccessful here, the pro-Essential Indexicality arguments elsewhere are still incomplete. Suppose, for example that the opacity of action-explanation contexts with respect to indexical expressions is cited as evidence for the special role of indexical thought in action. Suppose furthermore, that the source of opacity for indexicals is interestingly and importantly different from the source of opacity, e.g. names. There’s then a further question about whether these two supposed facts fit nicely together. If the specialness of opacity for indexicals derives from something unrelated to action, then even granting that specialness won’t rescue the pro-Essential Indexicality argument regarding action. Here’s an extreme way of making the point: suppose (implausibly) that all, and only, indexicals started with the letter “k,” and that there was a special problem in giving a theory of opacity for words starting with “k.” That fact would go exactly zero distance toward showing that the non-substitutivity of indexicals in action contexts showed something interestingly different from the non-substitutivity of names in those same contexts.

What we’ll refer to as “classic Frege cases” involve linguistic contexts in which the replacement of one of two co-referential terms with the other can result in a change in truth value. “Lois Lane believes that Superman can fly” is true. “Superman” is co-referential with “Clark Kent,” but if we substitute it for “Superman” we end up with a sentence that expresses something false. The Fregean line of thought then continues: since the two sentences have different truth values, they must express different contents. Since they express different contents, there must be a level of content over and above reference. Substitution failures of this kind occur not just in belief reports, but in reports of, roughly speaking, attitudes in general (hopes, fears, thoughts, doubts, etc.), and contexts that are parasitic on such reports. Just how widespread the phenomenon is, is an open question, one we will not take a stand on here.

There’s no consensus about how to explain such substitution failures. A range of theories has been proposed over the last century, and we won’t even attempt to give an overview. Suffice it to say that this is very difficult terrain and the explanation of opacity will fundamentally affect one’s overall theory of semantics, communication, and representation. Our goal in what follows is to consider whether substitution
failures involving indexicals present problems of a distinctive kind. We first present six of the classic puzzle cases (or thought experiments or examples) involving indexicals. For each of them we show that they pattern in obvious ways with classic Frege puzzle cases.

4.1. Puzzle Cases Involving Indexicals With Frege Counterparts

We repeat here some famous cases used to illustrate the behavior of indexicals in opaque contexts and for each we give what we’ll call their “Frege counterpart.” A Frege counterpart is a case like the original one in all relevant respects except that the substitution failure involves names instead of indexicals.

Case 1: The Messy Shopper

We start by reminding the reader of such a pair from Chapter 3. Perry’s messy shopper goes as follows:

I once followed a trail of sugar on a supermarket floor, pushing my cart down the aisle on one side of a tall counter and back the aisle on the other, seeking the shopper with the torn sack to tell him he was making a mess. With each trip around the counter, the trail became thicker. But I seemed unable to catch up. Finally it dawned on me. I was the shopper I was trying to catch. I believed at the outset that the shopper with the torn sack was making a mess. And I was right. But I didn’t believe that I was making a mess. That seems to be something I came to believe. And when I came to believe that, I stopped following the trail around the corner, and rearranged the torn sack in my cart. (1979, p. 3)

In Chapter 3 we gave the following Frege counterpart:

Pushing my cart down the aisle I was looking for CK (Clark Kent) to tell him he was making a mess. I kept passing by Superman, but couldn’t find CK. Finally, I realized, Superman was CK. I believed at the outset that CK was making a mess. And I was right. But I didn’t believe that Superman was making a mess. That seems to be something that I came to believe. And when I came to believe that, I stopped looking around and I told Superman to clean up after himself. My change in beliefs seems to explain my change in behavior.
Case 2: The Lost Hiker

Here’s another famous case from Perry’s “The Essential Indexical,” this one involving the indexical “here”:

The author of the book, *Hiker’s Guide to the Desolation Wilderness*, stands in the wilderness before Gilmore Lake, looking at the Mt. Tallac trail as it leaves the lake and climbs the mountain. He desires to leave the wilderness. He believes that the best way out from Gilmore Lake is to follow the Mt. Tallac trail up the mountain to Cathedral Peaks trail, on to the Floating Island trail, emerging at Spring Creek Tract Road. But he does not move. He is lost. He isn’t sure whether he is standing beside Gilmore Lake, looking at Mt. Tallac, or beside Clyde Lake looking at Jack’s peak, or beside Eagle Lake looking at one of the Maggie peaks. Then he begins to move along the Mt. Tallac trail. If asked, he would have explained the crucial change in his beliefs this way: “I came to believe that this is the Mt. Tallac trail and that is Gilmore Lake.” (1979, p. 4)

In the following Frege counterpart, assume that Byrde Lake is identical to Gilmore Lake but the hiker doesn’t know that:

Our hiker is lost. He knows he’s standing next to Byrde Lake, but doesn’t know whether Byrde Lake is Gilmore Lake, Clyde Lake, or Eagle Lake. As a result, he isn’t sure whether he is standing beside Gilmore Lake looking at Mt. Tallac, or beside Clyde Lake looking at Jack’s peak, or beside Eagle Lake looking at one of the Maggie peaks. Then he begins to move along the Mt. Tallac trail. If asked, he would have explained the crucial change in his beliefs this way: “I came to believe that Gilmore Lake is Byrde Lake.”

Case 3: Lingens the Amnesiac

In some of Perry’s cases it’s slightly less transparent that we are faced with nothing but a familiar Frege puzzle case. Perry’s discussion of Lingens (which is picked up again by Lewis in his 1979) is an example:

1 Some people will think that this sentence reports a *de se* belief of the hikers, and thus worry about whether this is a genuine Frege counterpart. But it is the opacity in the object position (“Byrde Lake”) that is of interest here, not the opacity in the subject position. It is trivial to construct slight variations of the example which remove all *de se* elements.

2 What follows is a slightly simplified version of the case, which focuses on knowing “who he is” rather than “who he is and where he is” at the same time.
An amnesiac, Rudolf Lingens, is lost in the Stanford library. He reads a number of things in the library, including a biography of himself. He believes any Fregean thought you think might help him. He still won’t know who he is no matter how much knowledge he piles up, until that moment when he is ready to say, “I am Rudolf Lingens.”

About this case, first note that it raises a set of issues that have to do specifically with the peculiarities of the “knowing who” locution. What does it take to know who you are talking to? It’s tricky to give a stable answer to this question. It appears to depend in complex ways on what is a salient identifying property in context. So, if I’m talking to person P who happens to be the Queen of Sweden, knowing a lot of information of the form, “She’s the person with property F,” might not suffice unless I also know that the person I’m talking to is the Queen of Sweden. Analogously, if Lingens is the King of Sweden, there are contexts in which we wouldn’t say that he knows who he is unless he knows he is the King of Sweden. In part, this is what the famous example is playing on; by describing a scenario in which Lingens is an amnesiac and relying on the assumption that one thing amnesiacs aim to do is recall their names, the reply “I am Rudolf Lingens” is made salient (this is obviously an enormous simplification, but the basic idea should be clear enough).

If we replace the “know who” claim with a straightforward “know that” claim, what remains is this:

An amnesiac, Rudolf Lingens, is lost in the Stanford library. He reads a number of things in the library, including a biography of himself. He believes any Fregean thought you think might help him. He still won’t know that he is Lingens no matter how much knowledge he piles up, until that moment when he is ready to say, “I am Rudolf Lingens.”

Now it is easy to give a Frege counterpart:

We are trying to locate Superman. We are in the Stanford library, and read a number of things in the library, including a biography of Clark Kent. We believe any Fregean thought you think might help us. We still won’t know that Clark Kent is Superman no matter how much knowledge we pile up, until that moment when we are ready to say, “Superman is Clark Kent.”

Lingens knows that Lingens is Lingens. He does not know that he is Lingens, even though the referent of that occurrence of “he” is Lingens. No
matter what descriptive information he gets about Lingens, say, the F, he can wonder whether the F is Lingens.

4.2. Some Irrelevant Cases

Not all cases that are discussed in the literature on indexicals have obvious Frege counterparts, but those that don’t are not about opacity at all, and so are not in the business of establishing Motivational Distinctiveness.

Case 4: Hume in His Study

Perry says:

Let us imagine David Hume, alone in his study, on a particular afternoon in 1775, thinking to himself, “I wrote the Treatise.” Can anyone else apprehend the thought he apprehended by thinking this? First note that what he thinks is true. So no one could apprehend the same thought, unless they apprehended a true thought. Now suppose Heimson is a bit crazy, and thinks himself to be David Hume. Alone in his study, he says to himself, “I wrote the Treatise.” However much his inner life may, at that moment, resemble Hume’s on that afternoon in 1775, the fact remains: Hume was right, Heimson is wrong. Heimson cannot think the very same thought to himself that Hume thought to himself, by using the very same sentence. (Perry 1977, p. 487)

The reason Heimson cannot use the sentence “I wrote the Treatise” to think the very same thought that Hume thought by using that sentence is that the referent of “I” depends on the speaker/thinker. That entirely obvious point is the only one made in this example.³ (We should note however that the case can be developed in ways that would make a non-trivial point about indexicality. One natural way of doing so is to focus on the different behavioral consequences of Hume and Heimson entertaining the same content. We addressed that distinctiveness claim in Chapter 3 and won’t rehearse our arguments against it here.)

³ It is worth re-emphasizing that Perry moves here uncomfortably back and forth between talk of language and talk of thoughts. We refer the reader to our comments on the problems with this move in Ch. 2.
Case 5: The Department Meeting

Some of Perry’s other famous cases are (at least on one reading) equally obvious illustrations of what it is to have a variable character:

Suppose the department meeting is scheduled for noon, September 15, 1976. Then only at that time could we say something true with “The meeting starts now.” Now consider any of the informationally equivalent thoughts we might have had the day before, for example “The meeting starts at noon, September 15, 1976.” It seems that one could accept this the day before, and continue to accept it right through the meeting, without ever accepting [the first thought], and even rejecting it firmly precisely at noon, simply by completely losing track of time. (Perry 1977, p. 491)

Since “now” has a variable character, it follows trivially that one’s cognitive attitude to sentences containing it will differ from one’s cognitive attitude to sentences containing stable denotations for points in time. In general, suppose IND is an indexical that in a context C denotes an object o and that T is a non-indexical expression that denotes o in all contexts. It is certainly possible for someone to endorse the sentence “T is F” in C without endorsing an utterance of “IND is F” in C, since nothing guarantees that the agent is aware that IND denotes o in C. Speakers are not omniscient about the denotation of indexical expressions relative to contexts. This does not reveal any distinctive kind of opacity. 4

That said, it is worth separating out two aspects of this example. There’s the “only at noon can you say something true by saying ‘The meeting is now’” bit. That, we have just argued, is a completely trivially corollary of having a variable character. Then there’s Perry’s claim that you could accept “The meeting starts at t” all the time, even when rejecting “The meeting starts now” at noon. That does have straightforward Frege counterparts. You can accept “Superman is dangerous” all the time, even when rejecting “Clark Kent is dangerous,” and even when confronted with Clark Kent. So the distribution of acceptance data is just a standard opacity feature.

4 Andy Egan pointed out to us that it might be better to make this point by saying that, even if they’ve got complete knowledge of character, and so they know full well (though tacitly) that IND denotes o in c, they don’t know that they’re in c.
4.3. Two Confusing Cases: Prior’s “Thank Goodness That’s Over” and Lewis’s Two Gods

*Case 6: Prior on “Thank goodness that’s over”*

Here’s an influential passage from Prior’s paper, “Thank goodness that’s over”:

One says, e.g. “Thank goodness that’s over!”, and not only is this, when said, quite clear without any date appended, but it says something which it is impossible that any use of a tenseless copula with a date should convey. It certainly doesn’t mean the same as, e.g. “Thank goodness the date of the conclusion of that thing is Friday, June 15, 1954,” even if it be said then. (Nor, for that matter, does it mean “Thank goodness the conclusion of that thing is contemporaneous with this utterance.” Why should anyone thank goodness for that?) (1959, p. 17)

A lot of things are going on in that part of Prior’s paper. Ted Sider proposes the following interpretation of the argument:

i) If the B-theory is true, then when I say “Thank goodness that’s over,” the object of my relief is a B-fact

ii) If the object of my relief is a B-fact, then it would have been just as reasonable for me to thank goodness before the ordeal was over

iii) It’s not the case that it would have been just as reasonable for me to thank goodness before the ordeal was over

iv) Therefore, the B-theory is false. (2005, p. 2)⁵

This reading of the passage is interesting and maybe something like this is what Prior had in mind. Again, whether it’s a good argument against B-theories is beside the point in this context. We are now concerned with whether evidence is provided for Motivational Distinctiveness. On the Sider-reading of Prior’s passage, there’s no attempt to establish that. The argument, if it succeeds, provides evidence that attitudes like relief don’t attach to B-facts, but that’s not even the beginning of an

---

⁵ Sider does not endorse this argument—see pp. 20–21 of Sider 2001.
argument for a distinctive kind of opacity arising in connection with indexicals.\(^6\)

That said it is easy to reconstruct the Prior case so it has a straightforward Frege counterpart. Suppose now = time \(t\). Then (1) and (2) get different truth values, because “now” and “\(t\)” aren’t intersubstitutable:

1) I am relieved that the event is over now.
2) I am relieved that the event is over at time \(t\).

Put that way, the Frege counterparts are easy to produce:

3) I am relieved Superman is here.
4) I am relieved Clark Kent is here.

**Case 7: Lewis’s Two Gods**

Our final case is from David Lewis. In his paper “Attitudes De Dicto and De Se,” Lewis asks us to consider the following possibility:

Consider the case of the two gods. They inhabit a certain possible world, and they know exactly which world it is. Therefore they know every proposition that is true at their world. Insofar as knowledge is a propositional attitude,

\(^6\) While it’s not even the beginning of an argument for opacity, it may be the beginning of a different and potentially interesting argument. One could also hold that there is a difference in truth value between “I am relieved that the event is over now” and “I am relieved that the event is over at time \(t\)” because the two expressions “that the event is over now” and “that the event is over at time \(t\)” are not co-referential, but pick out different facts—with the first expression picking out an A-fact and the second expression picking out a B-fact. When read that way, the case can’t support any thesis about opacity, since it is not an instance of failure of substitutivity of co-referential expressions. But it could support a metaphysical thesis that there is a difference in the world between “perspectival” facts (like A-facts) and perspective-independent facts (like B-facts). We leave this metaphysical defense of Essential Indexicality largely outside the scope of this work, but make two brief comments. First, we suspect that in fact this is a case of opacity, and that the metaphysical thesis is tempting only insofar as one presupposes that the relevant contexts are extensional, and hence (erroneously) takes failure of substitutivity as evidence for non-co-reference. Second, even if there is a metaphysical distinction between A-facts and B-facts, it is far from obvious how to connect that metaphysical distinction to the phenomenon of indexicality that will make it relevant to our main thesis. (It is clearly possible, for example, to pick out A-facts non-indexically, or B-facts indexically. A-facts can be named, and B-facts can be demonstrated.) There is an extensive literature on the metaphysics of time. It is, unfortunately, impossible for us to engage with that entire literature in this book. Those who base their endorsement of an essential indexicality thesis on arguments based on the metaphysics of time will find little to engage directly with in what follows. Thanks to Andy Egan for pressing this issue.
they are omniscient. Still I can imagine them to suffer ignorance: neither one knows which of the two he is. They are not exactly alike. One lives on top of the tallest mountain and throws down manna; the other lives on top of the coldest mountain and throws down thunderbolts. Neither one knows whether he lives on the tallest mountain or on the coldest mountain; nor whether he throws manna or thunderbolts. Surely their predicament is possible. (The trouble might perhaps be that they have an equally perfect view of every part of their world, and hence cannot identify the perspectives from which they view it.) (1979, p. 520)

There’s a lot to say about this case—so much that we devote all of Chapter 5 to a discussion of it and related themes from Lewis. The central point to note in this context is that whatever is shown by Lewis’s thought experiment of the two gods, it is not about a distinctive kind of opacity that arises in connection with indexicality. On the face of it, Lewis is making a simple point in three steps. First, he individuates propositions as sets of possible worlds—i.e. in the most coarse-grained way propositions can be individuated. It’s a conception according to which the proposition expressed by “2+2=4” is the same as the proposition expressed by “3+3=6”; they both express the set of all worlds. Second, he notes that terms like “knows” are sensitive to more fine-grained distinctions in contents: Lois can know that Superman flies without knowing that Clark Kent flies even though the set of worlds picked out by that Clark Kent flies is the same as the set picked out by that Superman flies. Similarly, the god throwing thunderbolts can know what’s expressed by “The guy throwing thunderbolts lives on top of the coldest mountain” without knowing what’s expressed by “I’m the guy living on the top of the coldest mountain,” even though the two sentences express the same set of worlds. Lewis is pointing out something we all know: contents are more fine-grained than sets of worlds. This isn’t an uninteresting point, but it doesn’t and isn’t meant to establish Motivational Distinctiveness.

We take it that the ease with which Frege counterparts can be generated makes at least a prima facie case that the Perry/Lewis/Prior-style cases simply are familiar substitution puzzles and that nothing new is brought out that distinguishes indexicals from other referring expressions with respect to opacity. Here, then, is our challenge to those who
find such cases powerful: *What, if anything, distinguishes these cases from Frege puzzle cases?* We suspect the reply will be something along the following lines:

*You are right that the cases involving indexicals instantiate the same pattern as Frege puzzle cases. What makes them distinctive is that the classic Fregean response to such puzzles, which appeals to descriptive content, senses, modes of presentation, or what-have-you, is particularly difficult to implement in this case. That’s what makes indexicals semantically distinctive.*

Our reply will be centered on two claims:

**Claim 1.** Fregeans have a problem with indexicality, but it has nothing to do with opacity.

**Claim 2.** Perry’s famous arguments to the effect that indexicals present a problem for Fregeans have nothing specifically to do with indexicals—they are just Kripke’s objections to any form of Fregeanism.

Before we get to that, we’ll go over two preliminaries.

### 4.4. Preliminaries 1: Scope Restriction and Why Non-Fregeans Can Skip the Rest of This Chapter

As noted earlier, Fregean Diagnostic Distinctiveness is an instance of a more general Distinctiveness thesis, Diagnostic Distinctiveness, according to which:

*There might be some solution to the problem of opacity in which indexicals are treated in a very special way and that special treatment has enormous implications for how we should think about meaning and content more generally.*

An evaluation of the more general thesis would require evaluating every possible account of opacity—a task that goes far beyond what we can hope to achieve in one work. So we don’t claim to have refuted the more general thesis. We have chosen to focus on how Fregeans account for indexicality for three reasons. First, we hope that what we say about Fregean Diagnostic Distinctiveness will throw light on
the more general thesis of Diagnostic Distinctiveness. Second, the Perry-Lewis tradition has been Frege-centric from its very beginning. Two of Perry’s most discussed papers, “Frege on Demonstratives” and “The Problem of the Essential Indexical” both contain explicit discussion of Frege (and considers ways Fregeans can deal with indexicals). Third, and most important, the non-Fregean traditions of accounting for opacity as a matter of historical fact don’t make a big deal about indexicals. To get a sense of what we have in mind, here are four of the most influential non-Fregean traditions:

- So-called “Naïve Russellianism” (versions of which can be found, e.g. in Braun 1998, 2000, 2001a, 2001b; Richard 1983, 1987; Salmon 1986, 1989, 1995a; Soames 1987, 1989, 1995; and Tye 1978);
- Mark Richard’s view of Russellian annotated matrixes (see Richard 1990);
- Larson and Ludlow’s (1993) proposal that the contents of attitude reports are treated as Interpreted Logical Forms; and
- Donald Davidson’s (1968) paratactic theory.

It is not that these theories don’t talk about indexicals or make an effort to incorporate them. What we want to emphasize is this: insofar as these theories pay attention to indexicals it is in order to recognize their variable characters. Beyond that, indexicality plays no important role and certainly does not motivate any revisionary account of content. In short, no one can read these theorists and say: Look, indexicality is of central importance in their account of opacity and again in their account of what meaning is, so a theory of indexicality must have wide-reaching philosophical implications.

Before moving on, we briefly note a corollary of these points: if you are not a follower of Gottlob Frege, feel free to skip the rest of this chapter. For non-Fregeans it suffices to recognize the conclusion of the first part of this chapter: the examples standardly paraded in the Perry-Lewis tradition involving indexicals in attitude contexts are instances of familiar Frege puzzles and provide no reason to think indexicality is distinctive in this respect.

7 The point we make here is a simple textual point about the non-Fregean traditions. For a good overview, see McKay and Nelson (2010). In it, “the problem of indexicals” is discussed as a problem for Fregeans, but no mention of indexicals is made in connection with non-Fregean theories.
4.5. Preliminaries 2: Robust Fregean Senses vs. Descriptive Fregean Senses

One more preliminary: anyone familiar with the vast literature on Gottlob Frege’s work will recognize that it is difficult to discuss even Fregean Diagnostic Distinctiveness in full generality—there are simply too many versions of Fregeanism. In particular, there’s little agreement among Fregeans about what senses are. For our purposes, it’s helpful to divide Fregeans into two camps. There’s what we can think of as the “simple descriptivism” Kripke argued against in Naming and Necessity (Lecture I), where the sense of a referring expression can be given by a descriptive clause along the lines of “the F.” Other Fregeans think that what is needed is a more robust account of senses. They suggest thinking of senses, e.g. as narrow psychological roles (Loar 1988, Fodor 1989 and 1991, Block 1991, Chalmers 2003) or in epistemic terms as a sense having to do with the channels of information that a term is tied up with (e.g. Evans 1982). If you’re thinking about sense in one of these “extrinsic” ways, where the theory of sense co-opts resources from another area, then the burden of distinctive indexicality is just going to get deferred to the other area. So, if you think that senses are all about narrow psychological role, then you’re going to think that special problems or features of indexical senses are just going to be surface manifestations of underlying special problems or features of indexicality in action, in which case see Chapter 3. Or if you think that senses are all about epistemic management, then you’re going to think that special problems or features of indexical senses are just manifestations of special problems or features of indexicality in epistemology, in which case see Chapter 7. So the general formula here is, if you think that sense is something fancier than simple descriptive Fregeanism, then probably what you’re really interested in is one of the later or earlier chapters.

With these preliminaries out of the way, we turn to the claim that what makes indexicals distinctive is that the classic Fregean response to substitution puzzles, which appeals to descriptive content (senses, modes of presentation, or what-have-you) is particularly difficult to implement in this case. Our response is twofold and spelled out as Claims 1 and 2 in Sections 4.6–4.7.
4.6. Claim 1: Fregeans Have a Problem With Indexicality, But It Has Nothing to Do With Opacity

Followers of Frege tend to treat as a fundamental tenet that sense determines reference: same sense, same reference. From that it follows trivially that indexicals don’t have the same sense: different uses of “I” have different referents, so they can’t have the same sense. The underlying tension is between (i)–(iii):

i) Sense determines reference.
ii) Reference shifts with context.
iii) Sense doesn’t shift with context.

According to (iii), “I” has a sense that is stable between speakers. By (i), that sense fixes reference. It then follows that “I” can’t change its referent. But it does, as (ii) points out. So we can’t have (i)–(iii) together. Here’s a famous passage from Frege where the tension between (i)–(iii) is the underlying problem:

If someone wants to say the same today as he expressed yesterday using the word “today”, he must replace this word by “yesterday”. Although the thought is the same, the verbal expression must be different so that the sense, which would otherwise be affected by the differing times of utterance, is readjusted. The case is the same with words like “here” and “there”. In all such cases the mere wording, as it is given in writing, is not the complete expression of the thought, but the knowledge of certain accompanying conditions of utterance, which are used as means of expressing the thought, are needed for its correct apprehension. The pointing of fingers, hand movements, glances may belong here too. The same utterance containing the word “I” will express different thoughts in the mouths of different men, of which some may be true, others false. (1956, p. 296)

Putting aside additional problems that arise due to the endorsement of Frege’s view that the sense of a sentence is the thought it expresses, the lesson to learn from the tension between (i)–(iii) is simply that if you want a theory of meaning and reference, you have to take context-sensitive expressions into account. There are many ways to go here.
Evans accepts the view that different uses of indexicals don’t have the same senses.\(^8\)

Obviously, if a Fregean approach to this utterance is to be sustained, the demonstrative in context must have a sense, and a different sense in different contexts. (1985, p. 292)

Other Fregeans, such as Katz (2004), reject thesis (i), and endorse a view of sense that is insufficient to determine reference. What is not an option is to reject (ii).

This book is not devoted to Frege interpretation and this chapter is not an overview of recent exploration of Fregeanism. So we will not dwell much on the strength and weakness of the various responses here. We just want to make one important point: the problem Fregeans here face isn’t about the interaction between indexicality and opacity. It is a problem that arises even in languages without opaque contexts. So it provides no evidence for what we have called Fregean Diagnostic Distinctiveness.

### 4.7. Claim 2: Perry’s Famous Anti-Fregean Arguments Have Nothing Specifically to Do With Indexicality

Perry (1977) argues that indexicals present an obstacle for Fregean theories of meaning and attitude reports. The goal of this section is not to show that Perry’s arguments fail, but that they could have been run using names, rather than indexicals. The objections he raises are powerful, but have nothing specifically to do with indexicality.

\(^8\) Note that those who endorse Evans’s position should stop using some of the ideology currently popular in the literature on the \textit{de se}. That literature relies heavily on terminology such as “the first person way of thinking,” “the first person concept,” and “the \textit{de se}.” The use of the definite article here presupposes that there’s a unique first-person way of thinking (or concept). That’s what Evans gives up. If different occurrences of “I am \textit{F}” have different senses, then there is no one thing that is “the first-person thought that one is \textit{F}.” And if there is no one thing that is the first-person thought, then arguments that we need indexical contents to account for patterns of sameness in, e.g. action can’t be used.
Perry (1979) says that one of his goals is to undermine a Fregean view according to which thinking about an object, \( o \), using a referential term requires thinking of \( o \) under a concept. The Fregean view Perry attacks is described as follows:

Just before I straightened up the sack I must have come to believe some propositions with the structure \( a \) is making a mess, where \( a \) is some concept which I alone “fit” (to pick a phrase neutral among the different notions of a concept). When I say “I believe I am making a mess,” my hearers know that I believe some such proposition of this form; which one in particular is not important for the purposes at hand. (1979, p. 7)

Perry’s reply to this is familiar by now:

If that \( a \) is making a mess is what I came to believe, then “I came to believe that \( A \) is making a mess”, where \( A \) expressed \( a \), should be an even better explanation than the original, where I used “I” as a communicative shortcut. But, as we saw, any such explanation will be defective, working only on the assumption that I believed that I was \( a \). (1979, p. 7)

Perry points out that no Fregean substitution for “I” would give us a proposition with the same cognitive, motivational, and semantic effects as “I.” No matter what the proposed substitution \( S \) is, it is only on the assumption that Perry also has the belief he would express by “I am \( S \)” that these effects are preserved.

In a related passage, Perry emphasizes a different point. About the Fregean view that a conceptual ingredient is needed, he says:

...in order to provide me with an appropriate proposition as the object of belief, the missing conceptual ingredient will have to fit me. Suppose I was thinking of myself in the way described, but that I wasn’t bearded and wasn’t in a Safeway store—I had forgotten that I had shaved and gone to the A & P instead. Then the proposition supplied by this strategy would be false, while what I came to believe, that I was making a mess, was true....This strategy [the Fregean strategy] assumes that whenever I have a belief I would characterize by using a sentence with an indexical \( d \),

I believe that...\( d \)...
that there is some conceptual ingredient \( c \), such that it is also true that,

I believe that \( d \) is \( c \)
and that, on this second point, I am right. But there is no reason to believe this would always be so. Each time I say “I believe it is now time to rake the
leaves,” I need not have some concept that uniquely fits the time at which I speak. (1979, p. 8)

For those with even a slight familiarity with the literature, it should be clear that these arguments against the Fregean account of indexical reference are simple instantiations of familiar and general objections to Fregeanism found in Kripke. In particular, they are instantiations of the objections sometimes referred to as the “Problem of Misdescription,” the “Problem of Ignorance,” and the “Epistemic Argument.” We consider these in turn:

- **Comparison with Kripke’s Misdescription and Ignorance Problems:** Suppose you believe incorrectly Einstein was a great logician who proved the incompleteness theorems. If you say, “Einstein was a logician,” you still use “Einstein” to refer to Einstein, and say something false about him. You don’t refer to whoever satisfies the descriptive content that you associated with the name “Einstein” (and so say something true about Gödel). Note that Perry’s argument is an instance of the Misdescription Problem: Perry asks us to suppose that he believes incorrectly that he is the unique $D$, and hence use $D$ as the sense of the word “I.” Suppose Perry is happy, but the person who is in fact the unique $D$ is not happy. Perry utters, “I am happy.” What he says is true, but what is predicted as the content of his utterance is false. A related problem is sometimes called “the Problem of Ignorance.” As Kripke points out, we can use a name, “NN” to refer to NN, without being in possession of any information that uniquely identifies the referent of “NN.” The same point applies to indexicals: suppose Perry doesn’t take himself to have any uniquely identifying information about himself. (Maybe this is more plausible with a different indexical, like “today”: he doesn’t take himself to have any uniquely identifying information about this day.) Then there is no descriptive sense that he can associate with the word “I” as he uses it.

- **Comparison with Kripke’s Epistemic Argument:** Another pillar of *Naming and Necessity* is the “Epistemic Argument.” Suppose the descriptive content I associated with “NN” is “the $F$” Getting to know the information expressed by “NN is the $F$” requires work that’s not required to come to know that “The $F$ is the $F$”
The latter is arguably analytic, maybe also a priori. Suppose he thinks that the shopper he is following is making a mess, but is not yet taking action to check his own groceries. Suppose, for any given description $D$, he comes to believe that the unique $D$ is making a mess, but does not believe that he is the unique $D$. Then, presumably, he will still not check his own groceries. So explanation of action—the relevant shift in cognitive state—requires in addition the belief that he is the unique $D$. So $D$ cannot be giving the sense of “I” for the shopper, or this additional belief would not be needed. Put another way, for any given $D$, an agent can take it to be uncertain whether he is the unique $D$. But he can’t take it to be uncertain whether he is himself, or whether the unique $D$ is the unique $D$. So the problem of cognitive significance is a manifestation of the Epistemic Argument.

To show that what we have just presented is not an idiosyncratic reading of Perry, consider the following completely orthodox presentation of Perry’s arguments in David Braun’s entry on “Indexicals” in *The Stanford Encyclopedia of Philosophy*. Here is Braun’s summary of Perry’s objections to Fregeanism:

Suppose that Fred sincerely utters “Today is July 4, 2001” on July 3, 2001. According to Frege’s theory, there is some non-indexical description that captures the sense of Fred’s utterance of “today”. But, as Perry (1977) points out, Fred may find it very difficult to produce a description that contains no indexicals that he would be willing to substitute for his utterance of “today” and that uniquely picks out one day. (In fact, in the previous paragraph the definite descriptions that we substituted for Fred’s utterances of “you” and “he” almost certainly do not have unique referents.) (Braun 2012, §4.3)

This is an instance of Kripke’s “problem of ignorance.” Braun then goes on to attribute an instance of Kripke’s “misdescription argument” to Perry:

Moreover, Perry says, it is possible that any such description that Fred would provide would not refer to the same day as his utterance of “today”. For instance, Fred might be inclined to express the sense of his utterance of
“today” with the description “the day in 2001 during which Americans celebrate the signing of the Declaration of Independence”. If so, then Frege’s theory entails that Fred’s utterance of “today” on July 3, 2001 refers to July 4, 2001. So Frege’s theory entails that Fred’s utterance of “Today is July 4, 2001” is true, and that he asserts and believes a truth. But Fred’s utterance is false (no matter how he is inclined to describe the day on which he produces his utterance), and he asserts and believes a falsehood. So Frege’s theory is incorrect. (2012, §4.3)

Finally, Braun goes on to describe a related objection to Fregeanism from David Kaplan:

Kaplan raises a related modal problem for Frege’s view. Suppose that Fred utters “If I exist, then I am speaking”, and suppose that the sense of his utterance of “I” can be expressed by “the person speaking”. Then on Frege’s theory, Fred’s utterance expresses the same sense as “If the person speaking exists, then the person speaking is speaking”. This expresses a necessary truth, but Fred’s utterance clearly does not. (2012, §4.3)

Again, this patterns exactly with a central argument in Kripke against Fregeanism about names: if the sense of a name, “NN” can be expressed by a definite description, “the F,” then the sentence “NN is the F” will express a necessary truth (we have called this the “Epistemic Argument”, this version of it is sometimes called the “Modal Argument”).

4.8. Taking Stock

This chapter has made two fairly simple points.

a) The examples of indexicals not being substitutable with co-referential terms in opaque contexts pattern with, and are instances of, familiar Frege puzzles.

b) The diagnosis of why these substitutions fail is just the same as the diagnosis of why substitution of co-referential names fails in opaque contexts.

So both Motivational Distinctiveness and Diagnostic Distinctiveness are false.
4.9. Objection 1: You Haven’t Refuted Fregeanism

For all you have said, Fregeanism is true, and that theory has a problem with indexicals. Maybe it’s not a problem that arises specifically in connection with an account of opacity, but it’s a problem nonetheless. As you pointed out, the following three assumptions are incompatible:

i) Sense determines reference.

ii) Reference shifts with context.

iii) Sense doesn’t shift with context.

Isn’t it possible that in order to preserve Fregeanism (a desideratum you don’t rule out) we need to recognize a level of content that is irreducibly indexical? Isn’t it possible that an interesting thesis, recognizable as a version of Essential Indexicality, emerges from the attempt to fix Frege?

Reply. First a brief digression: if, like us, you are not a follower of Gottlob Frege, this whole issue is bizarre. We non-Fregeans say that a theory with commitments that makes it a mystery how there can be expressions with variable characters is hopeless and not even a worthy fixer-upper. Keep in mind that we are emphatically not trying to show that there are no facts about indexicality that create trouble for any imaginable semantic theory. That would obviously be stupid. What would matter to us is the existence of semantic theory with three features: that’s there’s good reason to think it is right, that it runs into trouble with indexicality, and that the required fixes provide support to a version of Essential Indexicality. So that’s a very thin window of opportunity to get through. If the semantic theory gets into trouble with indexicality, then that’s reason to doubt the claim that it’s the right semantic theory, so the first two criteria are hard to put together. That’s why it’s a perfectly natural reaction to the conflict between the Fregean theory of senses and the facts about contextual reference shift to think: so much the worse for Fregeanism. It’s not that one has to think that, but it’s always going to be an option on the table, and the position of Fregeanism as “definitely right” is going to take a hit when the tension with indexicality emerges.

However, arguing that—and more generally that all forms of Fregeanism should be rejected—goes beyond the scope of this book,
so here is a more direct response to those who have Fregean sympa-
thies: As a matter of fact, the attempts to fix Frege typically involve
nothing more than the recognition of variable characters (in Kaplan’s
sense). No new or philosophically interesting features are added except
by those who advocate what we called “robust” Fregean senses (see
Section 4.5) and robust Fregean senses are motivated by considerations
discussed (and rejected) in other chapters, so won’t be discussed here.
The main point we want to emphasize at this juncture is that even if
one does continue to think that Fregeanism is right when one notices
the conflict between (i)–(iii), what you are faced with is just an obvi-
ous technical problem in holding senses fixed, having senses determine
referents, and allowing references to shift. Any move is going to have to
drop one of those three—in fact, it’s going to have to drop (i) or (iii),
because (ii) just looks like a data point. So the claim is going to have
to be that there’s no way to drop (i) or (iii) while retaining what makes
Fregeanism Fregean (and attractive) without endorsing an interesting
version of Essential Indexicality. Even here, we’re not arguing that it’s
impossible that this situation emerge. We’re just saying that no one has
made the case that it does, and that the case doesn’t seem to be there
for any semantic theory that anyone has proposed.

As an illustration of the point just made, consider a version of
Fregeanism (at least the proponents call it that) which is at least a
candidate for one that could be described as incorporating a version
of Essential Indexicality. It is a view that can be thought of as reject-
ing assumption (iii). John McDowell (1977), Peter Ludlow (1999), Saul
Kripke (1980), Stephen Schiffer (1987), Gareth Evans (1982), and oth-
ers have all flirted with or endorsed the thought, in one version or
another, that Fregeans should use indexicals to give the sense of index-
cicals. We will focus here on the work of Peter Ludlow, whose views
trace back to McDowell and Evans. 9 At the core of Ludlow’s version
of Fregeanism is the proposal that indexicals be used on the right-hand
side of semantic axioms in a meaning theory. To see what he has in
mind, some background is needed first. John McDowell (1977) pro-
posed that (1) and (2) display the senses of the object language.

  9 Related ideas are also found in Heck 1995 and Rumfitt 1993.
1) “Cicero is bald” is true iff Cicero is bald.
2) “Cicero is bald” is true iff Tully is bald.

In some sense, these two axioms state the same truth conditions, but they show the difference in meaning between “Cicero” and “Tully.” As Ludlow puts it, “What one wants from an adequate T-theory is that it gives the truth conditions in such a way that the senses of the object-language sentences are correctly displayed in the metalanguage.”

Like others in this tradition, Ludlow distinguishes between (a)–(c):

a) What the truth conditions literally state,

b) The way in which the truth conditions are represented, and

c) The sense displayed by the truth conditions so represented.

Rumfitt (1993) and Ludlow (1999) extend this style of view to indexicals. Ludlow’s (3) illustrates this strategy:

3) “I have a meeting now” is true just in case I have a meeting now.

(Where we give the content at the moment it is being used.) Ludlow says that we can think of theorems such as (3) as displaying the senses of the indexical expressions on the left-hand side. Of course, this isn’t going to help us if we want to interpret an utterance of “I have a meeting now” made yesterday. So Ludlow adds what he calls a theory of the cognitive dynamics of using expressions to track temporal senses (see Ludlow 1999 and 2007). If that theory works out, we end up with something like (4):

4) An utterance by me yesterday of “I have a meeting now” meant that I had a meeting then (yesterday).

We think this strategy is problematic for several reasons, but we won’t go into those here. Two points are important for our current purposes. First, on this view, there’s nothing special or distinctive about

10 One problematic feature of the view is that it implies that the true semantics of language can only be stated in the right context. In particular, the full semantics of the language can’t be said at one time, or in one place, or by one person. Prima facie it is bizarre to be told the following: Let me tell you our best scientific theory of how English works. I’ll have to go stand on the other side of the room to tell you part of it. There are of course moves that can be made here (maybe a theory of cognitive dynamics works for moves between “yesterday” and “today,” but it is hard to see how to extend it to “here”: there’s no automatic replacement the way we have for “today”), but we won’t explore those moves here.
indexicals. After all, according to this view, you can’t give the sense of “Superman” in a “Superman”-free language, either. If you think that can be done by using a suitable synonym for “Superman,” it just pushes the question back to whether you can find a suitable non-indexical synonym for “I,” and that’s more or less the central issue here. So, in brief, if you want to be a disquotationalist, of course you’ll think that there’s something essential about indexicals, but that’s because you think there’s something essential about every word in the language. Second, the view doesn’t involve commitment to a theory of meaning that goes beyond basic Kaplanianism. It does involve commitment to the intelligibility and usefulness of the stating vs. displaying distinction, but again, that has nothing to do specifically with indexicality.

4.10. Objection 2: You Haven’t Solved the Problem of Opacity

Sometimes when we present this material, audiences have been frustrated by the argumentative strategy of this chapter. They say: Look, you can’t know that there’s no distinctive opacity generated by indexicals unless you tell us what your favorite theory of opacity is, and then defend it. As long as you’re neutral on how to deal with opacity, there’s always the possibility that Diagnostic Distinctiveness is correct. In short, you owe us an account of opacity.

Reply. It is true that we haven’t given (and won’t give) you our theory of opacity. If we did that it would a) take a very, very long time—it would in effect be a separate book project, and b) it would make our arguments hostage to the acceptability of that theory. We are convinced that we can address Diagnostic Distinctiveness by showing that the issues that arise are either theory-internal problems for followers of Gottlob Frege, have nothing to do specifically with indexicality, or are trivial corollaries of context sensitivity. Whatever your view of those issues, it will provide no support for the target of this book.

We hope it is crystal clear that we don’t deny that a theory of opacity might appeal to characters (we also don’t endorse this, since we don’t endorse any specific theory of opacity in this work).
It is a triviality that there are non-constant characters, and that they are distinct from constant characters. We are not denying that such characters might play a role in an account of some instances of opacity—a difference in character might create a difference in substitution behavior. What we do deny is that the category of non-constant character plays a distinctive role in the theory of opacity. Whatever it is that explains opacity, it is a feature that can be manifested by expressions of both constant and variable character. Talking about character, then, does nothing to illuminate opacity (even if it gives some local illumination about the substitution behavior of some particular term). As a result, opacity phenomena give us no reason to think that variable character has wider implications throughout philosophy. It is simply a matter of a local solution to a small feature of the (so-called) opacity problem. That’s certainly a little bit philosophically interesting—we don’t deny that. But if that’s all you think Essential Indexicality amounts to, you’re on our side—i.e. you agree that indexical thought and language plays no important explanatory role in an account of agency, immunity to error, perceptual content, information gathering, or more generally, in an understanding of perspectivity and the subjective–objective distinction. You also agree that indexicals don’t generate interesting new kinds of opacity puzzles (or any new insights into the phenomenon of opacity).

4.11. Objection 3: To Refute Essential Indexicality You Need a Theory of Content and to Have a Theory of Content You Need a Theory of Opacity

We want to end this chapter by addressing directly one concern in this vicinity raised by an anonymous reader for OUP. Doing so will help clarify our aims in this book. The reader says:

*It seems to me that the issues raised by indexicals and indexical attitudes are hard to cleanly separate from the broader question of what the correct theory of opacity is, which is closely connected to the question of what the correct theory of content is.*
The reader then asks how we can avoid committing to an account of opacity. More generally: how can we defend the view that indexicality is philosophically shallow without a full theory of indexical content (which won’t be forthcoming without an account of opacity)?

Reply. This is where it is important to keep in mind that our goal is modest. For all we say in this work, it is certainly possible that someone develop a theory of indexical content that has all kinds of deep philosophical significance. The question we address in this book is whether the literature that traces back to Perry and Lewis’s seminal papers provides reason to think that a theory of indexicality will play an important explanatory role in philosophy (where importance is spelled out as one of the EI-theses). The following two analogies might be helpful. First, suppose someone claims to have made the remarkable discovery that co-referential names of US Senators can’t be substituted *salva veritate* in attitude contexts. It turns out, this person tells us, that when a US Senator has two names, “NN” and “NN1,” it is possible for someone to believe that NN is F, but not believe that NN1 is F. In response to this alleged discovery we would say: okay, but that’s just an instance of something we already knew: it seems that in general co-referential terms aren’t substitutable *salva veritate* in attitude contexts. Now, suppose the discoverer of Senator-non-substitutivity then replies: Oh well, so you say, but how do you prove that names for US Senators aren’t special? In particular, how can you say that without committing yourself to what the correct theory of opacity is? In response to this one should say: Of course we can’t prove that, but it’s a reasonable research strategy to ignore that option at this stage since we have no positive reason to think names for US Senators are special in this way. More generally, we would want to dismiss this person by saying something like: Look, opacity is a great mystery that philosophers and linguists have struggled with for more than 100 years. To insist that we solve that problem before we reject US-Senatorial Name Distinctiveness would be completely unreasonable. We think that an insistence on indexicals being opacity-distinctive (and thus potentially giving us reason to revise our theory of content) is just as unmotivated as insistence on US-Senatorial Name Distinctiveness.

A second analogy: Imagine a complete and true theory of pure quotation (i.e. expressions like “and”). Question: Have the last thirty
years of literature on quotation given us reasons to think that a theory of quotation will play an important role in an understanding of a wide range of philosophical phenomena, e.g., agency, perception, epistemology, self-knowledge, etc.? The answer is “no,” though of course a theory of quotation is important in semantics (since no semantic theory is complete without it) and for understanding opacity (since quotational contexts are opaque). We’re arguing that indexicality is like that (except probably less important since quotation is, at least arguably, a distinctive kind of opaque context). The deflationary view of indexicality promoted in this work is important, we think, in the current philosophical climate, since an enormous amount of recent literature takes Perry and Lewis to have shown that a theory of indexicality will have significant philosophical implications (spelled out as the various EI-theses). Our goal is to undermine that assumption.

Yet another way to put these points: suppose we told you our view of opacity: we can substitute co-referential terms *salva veritate* in belief contexts. We explain away the appearance of opacity using some version of the kind of theory proposed, e.g., by Scott Soames (2002). You’ll no doubt be skeptical. Then we will go through fifteen–twenty rounds of replies and counter-replies, largely about issues that have nothing to do with indexicality. At the end of that you’ll most likely respect our position, but won’t be converted (since that hardly ever happens in philosophy). And so at that point we want to say: look, even if you’re not convinced, that’s not a good enough reason to endorse an Essential Indexicality thesis because internal to your own theory you don’t have sufficient reason to endorse such views. It is that last point that will matter at the end of the day, and so we didn’t need to go via the attempt to convert you to a hard-core direct reference theorist.

### 4.12. Brief Summary of This Chapter

This chapter has five central take-home messages:

1) Frege counterparts can be generated for all classical examples in the essential indexicality literature and this makes at least a prima facie case that the Perry/Lewis/Prior-style cases simply
are familiar substitution puzzles, which bring out nothing new that distinguishes indexicals from other referring expressions with respect to opacity.

2) This presents those who find the Perry–Lewis cases powerful with an important challenge: what, if anything, distinguishes these cases from Frege puzzle cases?

3) The response we expect to get to this challenge is that they are distinctive in that the classic Fregean response to such puzzles is particularly difficult to implement in these cases.

4) Our response to the response in (3) is twofold. First, the Fregean problem with indexicality has nothing to do with opacity—it is a problem of reconciling the Fregean commitments to the claims that sense determines reference and indexical senses are stable between contexts with the fact that indexicals change referents between contexts. Second, the anti-Fregean arguments found in Perry 1977 and 1979 are simply instantiations of Kripke’s objections to Fregeanism about names. The examples Perry uses happen to involve indexicals, but that’s an accidental feature of the cases. Exactly the same points could have been made using names.

5) In the final sections of this chapter we have made some important points about the overall strategy (and coherence) of the project we have set ourselves in this book. We showed how a deflationary project such as ours can be achieved without relying on a commitment to a particular theory of opacity (or a particular theory of what meaning is).
Chapter 5

Lewis on the De Se, Self-Ascription, and Centered Worlds

In his paper “Attitudes De Dicto and De Se” (1979) Lewis uses examples and data from Perry and Castañeda to motivate a radical revision of classical views about content and propositional attitudes. The classical views are those we have been assuming (and indirectly arguing for) throughout this work: to believe something is to believe a proposition; propositions are the bearers of truth values; and propositional attitudes are properly so-called because they are relations between thinkers and propositions. According to Lewis, Perry-style data provide reasons for rejecting the classical view and instead endorsing a view that appeals to self-ascriptions of properties. Lewis compares his view to Quine’s proposal that contents be construed as centered worlds and for complex historical reasons Lewis’s proposal is now often described as a centered-worlds view. The currently widespread acceptance of (variations of) Lewis’s proposal might be the most lasting legacy of the Perry–Lewis tradition.

This chapter has two central goals.

- First, we argue that Lewis offers no fundamentally new arguments in favor of his revisionary theory of content, and that the arguments that he does deploy, especially when taken as arguments directed at a theory of information (i.e. which attempt to show that in order to explain information gathering and processing we need a category of distinctive de se information) are systematically unconvincing.
Second, we argue that even if one does accept that there is a problem in the theory of information that needs solving, the revisionary theory of content itself does none of the work in solving the problem—if there is anything in Lewis’s view that solves any problem, it resides in the sparsely developed theory of self-ascription. This is what specifies the relation an agent needs to be in to a (Lewisian) content in order to be in a propositional attitude. Once one sees the role that self-ascription plays, it becomes clear that essentially the same technique could have been combined with standard theories of content to address Lewis’s worries.¹

5.1. Lewis on Propositions and Properties

We begin with an overview of Lewis’s critical and positive moves in “Attitudes De Dicto and De Se.” The centerpiece of Lewis’s critical argument is a consideration of two examples: Perry’s case of Lingens the amnesiac, and Lewis’s variation on that case, the case of the two gods. Lewis claims that “these examples suffice to establish my second thesis: sometimes property objects will do and propositional objects won’t” (p. 139), so we want to focus very carefully on how Lewis thinks the examples succeed in establishing this.² We’ll begin by considering in detail Lewis’s two gods case:

Consider the case of the two gods. They inhabit a certain possible world, and they know exactly which world it is. Therefore they know every proposition that is true at their world. Insofar as knowledge is a propositional attitude,

¹ Note that in pursuing this second goal we do something beyond what we have done in previous chapters (and what we will do in forthcoming chapters). We argue, in effect, that even if you disagree with us about the lack of motivation for Essential Indexicality, the theory proposed—in this case Lewis’s—fails in reaching its stated goal.

² The argument we extract from the two gods case is not the only argument Lewis gives in “Attitudes De Dicto and De Se,” although it will be our focus here. Lewis also considers the case of mad Heimson, who believes himself to be Hume. Lewis uses this case to run an argument from narrow content:

Heimson may have got his head into perfect match with Hume’s in every way that is relevant to what he believes. If nevertheless Heimson and Hume do not believe alike, then beliefs ain’t in the head! They depend partly on something else, so that if your head is in a certain state and you’re Hume you believe one thing, but if your head is in that same state and you’re Heimson you believe something else. Not good. (Lewis 1979, p. 142)
they are omniscient. Still I can imagine them to suffer ignorance: neither one knows which of the two he is. They are not exactly alike. One lives on top of the tallest mountain and throws down manna; the other lives on top of the coldest mountain and throws down thunderbolts. Neither one knows whether he lives on the tallest mountain or on the coldest mountain; nor whether he throws manna or thunderbolts. Surely their predicament is possible. (The trouble might perhaps be that they have an equally perfect view of every part of their world, and hence cannot identify the perspectives from which they view it.) (p. 139)

Let’s first consider a simple but powerful (as we’ll see, perhaps too powerful) argument that can be extracted from this case:

P1: Each god knows (ex hypothesi) every true proposition.

P2: Neither god knows which of the two gods he is, or which mountain he lives on.

P3 (tacit): Which god he is, or which mountain he lives on, is something for each god to know.

C1: Neither god knows everything.

C2: There is some knowledge that is not knowledge of a true proposition.

Since we don’t, and don’t see any reason to, endorse the claim that all content is narrow content, we reject a crucial premise of the argument, and don’t see it as a compelling motivation for a thesis of Essential Indexicality. We would also note that it’s far from obvious that Heimson can get his head into a perfect match with Hume’s while at the same time believing, unlike Hume, that Heimson is Hume. Certainly Lewis gives us no argument for this possibility. Consider an analogous case. Schmeimson believes that water is dry, unlike Hume, who believes it is wet. But Schmeimson has got his head into perfect match with Schume’s in every way that is relevant to what he believes! So to keep beliefs in the head, we need to remove the wet/dry distinction from content.

Maybe it’s impossible for Schmeimson to get his head into perfect match with Schume’s in every way that is relevant to belief, because by hypothesis Schmeimson and Schume believe different things (about water)? But then why is it not also impossible for Heimson to get his head into perfect match with Hume’s, as already noted? It begs the question, of course, to hold that Heimson does believe the same as Hume—that’s meant to be the conclusion. Or maybe it’s impossible for Schmeimson to get his head into perfect match with Schume’s because the distinction between wet and dry water beliefs, unlike the distinction between Heimson-being-Hume and Hume-being-Hume beliefs, has functional consequences on action. We’re skeptical that there’s a clear case for this, but in any case, this moves the argument onto the territory of the discussion of functionalism in Ch. 6. Thanks to Juhani Yli-Vakkuri for pressing us to address this argument from Lewis.
This argument would, of course, refute the traditional picture—it would show that there was a piece of information each god lacked, but a piece of information for which we have no account in the traditional account of propositions. But as it stands, the argument clearly begs the question. If we simply write into the two gods case that each god knows *every true proposition*, then a fan of the traditional picture will hold that the god on the tallest mountain a fortiori knows that he is the god on the tallest mountain, since the proposition *that he is the god on the tallest mountain* is among the true propositions. Any attempt to specify something not known by one of the gods will be taken as a specification of a proposition not known, and thus as a rejection of the omniscience premise, P1.

To keep the argument from being question-begging, Lewis couples it with a particular view of what propositions are. Providing a substantive theory of propositions allows the scope of premise P1 to be delimited, so that Lewis can then argue that P1 can be non-question-beggingly combined with the assumption of godly locative ignorance. Lewis’s particular view is, of course, that propositions are sets of possible worlds. Given this view, propositional omniscience amounts to knowledge of exactly which world is actual. Why is this sort of knowledge not enough for a god to know which god he is, or what mountain he is on? Lewis equates *knowing which world is actual* with *knowing which world one is in*:

What happens when he believes a proposition, say the proposition that cyanoacrylate glue dissolves in acetone? Answer: he locates himself in a region of logical space. There are worlds where cyanoacrylate dissolves in acetone and worlds where it doesn’t. He has a belief about himself: namely, that he inhabits one of the worlds where it does. (p. 518)

So propositional omniscience amounts to locating oneself up to the level of one-world accuracy. But the gods are in quest of considerably more accuracy than this: they wish to locate themselves up to the nearest mountain, not just the nearest world. As Lewis says about Lingens:

The more he reads, the more propositions he believes, and the more he is in a position to self-ascribe properties of inhabiting such-and-such a kind of world. But none of this, by itself, can guarantee that he knows where in the
world he is. He needs to locate himself not only in logical space but also in ordinary space. He needs to self-ascribe the property of being in aisle five, floor six, of Main Library, Stanford. (p. 520)

The two gods, says Lewis, need the same sort of thing that Lingens needs: knowledge that locates them more finely than to-the-nearest-world. Lewis’s suggestion is then that we take attitudes to have properties, rather than propositions, as contents. To stand in the belief relation to a property is to self-ascribe that property.

The solution to the peculiar cases of Lingens and the two gods is then generalized. All attitudes are relations to properties. What we used to take as propositional attitudes Lewis can now take to be attitudes toward a particular subclass of properties: the properties of occupying a given set of worlds:

Since I construe properties broadly, this thesis is not very bold. We have a one-one correspondence between all propositions and some properties. Whenever it would be right to assign a proposition as the object of an attitude, I shall simply assign the corresponding property. Since the correspondence is one-one, no information is lost and no surplus information is added. The attitude is equally well characterized either way. And since it is easy to go back and forth, there can be no significant difference in convenience. (p. 516)

We can more generally re-identify standard classes of attitudes within Lewis’s system:

• Attitudes de dicto (descriptive/general attitudes) are self-ascriptions of properties that do not discriminate within world boundaries.
• Attitudes de re (object-dependent attitudes) are self-ascriptions of relational properties (typically ones that cut across world boundaries), properties of being $R$-related, for some appropriate acquaintance relation $R$, to an object that has some further property.

3 Recall that for Lewis, a property is a set of possible individuals. This enables Lewis to say that in self-ascribing a property, one is identifying oneself as one of the members of that set. Put another way, propositions and properties for Lewis are sets of things (typically much bigger things in the case of propositions rather than properties), and Lewis can then construe having a belief (with either a proposition or a property as a content) as a way of constraining the options for where one is. See our discussion of seductive metaphors in Sect. 5.2 for more details.
• Attitudes \textit{de se} are self-ascriptions of properties that do discriminate within world boundaries and are not relational properties of being acquainted with some object with a further property.

One point we want to stress here is that as we are reading Lewis, he is offering primarily an \textit{informational} puzzle in favor of his view that attitude contents are properties. The problem raised by Lingens and especially the two gods is that there is a \textit{piece of information} they don’t have, and no account of what that information is on the possible-worlds account of content. While there are also strands of action-explanation-based arguments in Lewis (especially in the interaction with his functionalism, which will be a point of discussion in Chapter 6), we will focus on the informational issues here, and allow our discussion of action and indexicality from Chapter 3 to cover those other strands.

Lewis’s view that attitude contents are properties is certainly a \textit{revisionary} theory of content, but why is it an \textit{indexical} theory of content? Two aspects of the account combine to make it indexical. First, because contents are properties rather than propositions, they are not in themselves truth-condition-determining in the normal sense. On Lewis’s way of thinking about this: when a content is a set of worlds, a notion of truth at a world drops out automatically—content $C$ is true at world $w$ if and only if $w$ is an element of $C$. But when $C$ is a property-like content that cuts across the boundaries of worlds, we don’t get the notion of truth at a world. If $C$ is a content and $w$ a world such that some but not all parts of $w$ are elements of $C$, we are delivered no clear sense in which $C$ is or is not true at $w$. Second, the notion of truth is then recaptured indexically via the pragmatics, i.e. through Lewis’s picture that to hold an attitude involving a property-content $C$ is to self-ascribe that content.

Lewis’s version of indexical content thus differs from, for example, Fregean versions of indexical content by achieving indexicality by subtraction rather than by addition. Lewisian property contents aren’t indexical because they include some extra element (a \textit{first-person concept}, for example), but because they don’t include something that the standard account would include (the person to whom the property is attributed), and then allow that missing element to be provided in the indexicalized theory of truth. It is for this reason that Lewisian
contents have been a natural tool for truth relativists to reach for in formalizing their accounts. Thus Andy Egan, one of the most prominent contemporary relativists, puts it this way:

The sort of relativist account that I favor is one on which some of our utterances express a kind of proposition that determines a truth-value not just relative to a world, but relative to a <world, time, individual> triple. (Call these centered worlds, since they single out not just a world, but also a privileged position—a center—within that world.) Such a theory of content is relativist because it says that some of our utterances have contents which can, for two individuals x and y, simultaneously determine different truth-values relative to x and y, even though x and y are worldmates. (2007, p. 5)

(We will return to the connection between property contents and centered worlds later.) Many of the points that emerge in the dispute between relativists and contextualists appear also in assessment of Lewis’s revisionary theory of content. Because our goal here is to move beyond the contextualism/relativism dispute, we attempt as much as possible to set aside issues that are narrowly specific to that dispute and focus on the broader question of whether Lewis has motivated a specifically indexical theory of content. We now turn to making a sequence of critical points about Lewis’s use of the two gods case as an argument for an informational need for indexical contents.

5.2. Response 1 to the Two Gods: Don’t Be Seduced by Metaphors

The rhetoric that Lewis uses to illustrate his property theory of content is evocative and compelling. Once we think of propositions as sets of possible worlds, the totality of our information state can then be thought of as a set of propositions, and hence as a set of sets of worlds. It’s then natural to model that set of sets of worlds as the intersection of the sets of worlds given by the propositions one believes.  

4 This is to accept that one’s information state is closed under metaphysical consequence, but we were very close to accepting such a commitment in using the possible-worlds framework in the first place.
Improvement of one’s information state is then a narrowing of that intersectively given set. So there’s a natural terminus to the process of improvement—one’s information state has been made as good as it can (consistently) be when it contains but a single world. Assuming all of one’s beliefs are true, that single world will be the actual world. So it’s natural to describe the process of inquiry as the process of determining which world is the actual world.

Natural, but not inevitable. A crucial bit of philosophy has been smuggled in at this point, so it’s worth pausing to see what it is, and how it might be avoided. In saying that the goal of inquiry is to determine which world is the actual world, we have introduced an indexical element (“actual”) into the description of the process. Once this element is introduced, it tends to spread—we move on to saying that believing that $p$ is believing that the actual world is in the set of worlds that is the content of $p$. But we didn’t have to indexicalize. The goal of inquiry can be described non-indexically: it is to believe all true propositions. Or, if one is wedded to the worlds talk, the goal is to have an information state consisting of a single world, which is an element of every true proposition. One might think that this description merely hides the indexicality, because truth here is covertly truth at the actual world. But (following Cappelen and Hawthorne 2009) we would resist the thought that monadic truth needs to be thought of as a special case of indexicalized truth at a world.\(^5\)

The talk of determining which world is actual is harmless so long as one remembers that it’s dispensable. But it’s then easy to move from one indexical to another. Talk of determining which world is actual transitions into talk of determining which world $I$ am in—on the assumption that the actual world is the one $I$ occupy. Lewis’s modal realism helps with this transition—when individuals are world-bound, the transition from “actual world” to “world $I$ am in” is seamless. When they are not, it is not—if $I$ am in many worlds, the actual world is not simply the world $I$ am in.

\(^5\) The point here isn’t that we have to have a theory of monadic truth—that’s a much bigger battle than we want to fight in this context. We just want to flag the ways in which Lewis’s presentation of the “propositions as possible worlds” framework makes non-mandatory choices about how to characterize the issues, and track the ways in which those non-mandatory choices shape the final theoretical stance that emerges in Lewis.
But now propositional omniscience has taken on a locative sense—to know everything is to know quite a bit about where I am. It’s to know where I am up to the nearest world. Once things are given that locative cast, though, it’s easy to think that there’s quite a bit of locative information still missing. Worlds, after all, are (often) quite big things. To know what world I am in is to know a lot about where I am, but it still leaves a lot of questions unanswered. At this point, we are primed for Lewisian semantics: we need contents that will tell us more precisely where we are. Since property contents in the limiting case contain exactly one object (the believing agent, when all beliefs are true), they allow full locative specification.

The slide from information-as-sets-of-worlds to information-as-indexical-identification, and from there to information-as-self-location, can feel inevitable under the weight of the Lewisian metaphors. But the metaphors are not arguments, and if one refuses to cast inquiry as an indexically given project, the grip of the metaphors evaporates.

And the metaphors cannot be taken at face value, and cannot non-question-beggingly drive the argument. (We should add that we don’t take Lewis to be using the metaphors as arguments. We take the central Lewisian argument to be the argument from the two gods case given in Section 5.1. The metaphors are rhetoric, and our goal here is to clear the air of the rhetoric so that the argument proper can be clearly evaluated.) A picture on which the goal of inquiry is to discover where one is, is peculiar at least twice over. First, it makes inquiry a strangely egocentric process. On this picture, everything one believes is about oneself. One tracks the world entirely by tracking how the world relates to oneself. Second, it makes information about oneself strangely hard to come by. Suppose you want to know where Paris is. Knowing that Paris is in France is surely a good way to know, at least roughly, where Paris is. But that Paris is France has as its content a set of worlds, so if the locative metaphor were consistent, we should think that in knowing that Paris is in France, we pin down the location of Paris only to within a (rather large) set of worlds. But knowing that Paris is in France gives you the location of Paris to within a few hundred miles, not to within billions of light years. So knowing where Paris is is crucially different (on the Lewisian metaphor) from knowing where you are. You can know where Paris is with much more precision than...
“somewhere in my set of open possibilities,” but apparently not where you yourself are. That can’t genuinely be right, so the metaphors can’t be taken at face value.

5.3. Response 2 to the Two Gods: Disentangling Propositions vs. Properties from De Dicto vs. De Re

Why does Lewis have two gods in his central case? The structure of the argument as given in Section 5.1 can be run with a single god: the single god knows all possible-worlds propositions, and hence knows exactly which world is actual, but does not know where he is in that world, and hence there is something he knows that is not a proposition. How does a second god add to the case? The obvious thought is that the second god somehow stops the first god from knowing which god he is (or which mountain he is on), by adding a confounding candidate. The same idea shows up in a similar case that Lewis gives in On the Plurality of Worlds:

Imagine someone who is completely opinionated, down to the last detail, about what sort of world he lives in and what goes on there. He lacks no belief about the world. For him, only one world is doxastically accessible…. And yet there may be questions on which he has no opinion. For instance he may think he lives in a world of one-way eternal recurrence, with a beginning but no end, with a certain course of history repeated exactly in every epoch; and he may have no idea which epoch he himself lives in. Every epoch of the world he takes to be his contains someone who might, for all he believes, be himself. He has no idea which one of them he is. If he did, for instance if he somehow became persuaded that he lived in the seventeenth epoch, he would believe more than he does. But he would not believe more about the world. The added belief would be not about the world, but about his own place therein. (1986, p. 28)

Again we can ask: why the additional complexity of the eternal recurrence? Given Lewis’s picturesque metaphor, our opinionated someone has, via his opinions, narrowed his location down to a world, but no further. Every resident of that world should be an eligible candidate for
Opinionated Man; there’s no reason Opinionated Man should restrict his options to his epochal copies.

We can draw two lessons from the apparently unnecessary gods and recurring epochs. First, Lewis’s argument depends on the plausibility of the thought that, despite propositional omniscience, there remains something that the god, or Opinionated Man, doesn’t know. There’s no direct argument for the existence of something unknown; we just have to be convinced by the example that the propositional knowledge doesn’t cover everything there is to be covered. Suppose Opinionated Man lives in a recurrence-free world, and is quite distinctive in that world—he is, for example, the only one in that world who always wears a purple hat. We assume Opinionated Man knows all propositions. What grip remains to the thought that he doesn’t know who he is? If the grip is diminished, that diminution casts doubt on Lewis’s picture of knowledge as self-location, because according to that picture, Opinionated Man’s ignorance is the same in both cases.

Second, Lewis’s introduction of the second god, and use of eternal recurrence in the case of Opinionated Man, suggests that he wants qualitative/descriptive identity between candidates for identity. This in turn suggests that qualitative/descriptive information can locate objects “in the normal way,” by providing location within a given world, rather than just narrowing down which world is actual. It’s tempting to think that the picture is this: Lewis begins with the idea that all information is qualitative/descriptive, and hence that all belief is de dicto. The inadequacy of this picture is brought out through eternal recurrence-type examples, which show that we can have information and beliefs about objects despite the lack of uniquely identifying descriptions. (So far we are following one of the argumentative paths of Naming and Necessity against descriptive Fregeanism.) In order to pick out one particular instance of qualitative duplicates, we need a referential anchor—some point that is given non-descriptively, from which we can then descriptively bridge to other objects.

Lewis, in his typical minimalist way, provides us with a single referential anchor. His treatment of de se beliefs using properties as semantic values is a way of providing a single exception to the de dicto nature of belief—one’s belief is about oneself because of the way in which truth conditions are tied directly to the possessor of the belief, not because...
one descriptively identifies oneself. The thought that some contents have their truth conditions linked to objects not in virtue of descriptive identification of those objects is, of course, a familiar one—it is just the idea of object-dependent contents. So it becomes natural to wonder why we should accept Lewis’s *de re* minimalism, on which it is only the self that is genuinely picked out *de re.* Even those who work within Lewis’s framework often extend Lewis’s *de re* monomania, allowing additional objects to be picked out directly in the truth conditions (see Ninan 2010).

A shift from talk of sentential semantic values as properties to talk of sentential semantic values as sets of *centered worlds* aids the introduction of these additional objects. Although Lewis’s preferred formulation is in terms of properties, he notes that his view can also be stated using the idea of centered worlds, as considered (and then rejected) by Quine. Quine talks about a cat that wants to be on the roof and fears being in the dog’s jaws. How do we model the cat’s mental state? Here’s Quine’s proposal as described by Lewis:

[Quine] therefore suggests that we take *centered possible worlds*—in effect, pairs of a world and a designated space-time point therein—and regard the wanted or feared state of affairs as a class of centered worlds. A centered world is centered on a cat therein if and only if the designated point is in the midst of the cat—more precisely, is the center of gravity of the cat’s pineal gland. The cat wants a class of centered worlds that are centered on a cat safely on a roof. He fears a class of centered worlds that are centered instead on a cat in the jaws of a dog. No centered world belongs to both classes. A problematic world with many similar cats is a world that belongs to the wanted class under some centerings and to the feared class under other centerings. (1979, p. 531)

This proposal, Lewis claims, is equivalent to the proposal that propositions be replaced by properties.

A class of centered worlds corresponds to a property. Most directly it corresponds to a property of space–time points, but also it corresponds to a property of cats. Let $X$ be a class of centered worlds; there corresponds to it the property

---

6 “Genuinely” in the sense in which proponents of the *de re* normally think about it. Lewis, of course, has his own account of the *de re* that is not *de se*, but on his account, such beliefs are descriptive aside from their ubiquitous *de se* component.
of being a cat on which some member of $X$ is centered. Let $Y$ be a property; there corresponds to it the class of exactly those centered worlds that are centered on a cat having the property $Y$. (Here I assume that one centered world cannot be centered on two different cats, cats who occupy the same place at the same time. To avoid that assumption, as perhaps we should, we might redefine centered worlds as pairs of a world and a designated inhabitant thereof.) By centering the worlds, Quine has in effect replaced propositions by properties as objects of the attitudes. (1979, p. 532, bolded emphasis added)

The contemporary versions of centered-worlds semantics typically follow Lewis's parenthetical remark and treat that as Lewis's view. They then add additional centers, such as time. Here is a fair summary of the contemporary orthodoxy by Ninan:

Lewis's idea is that we should represent attitudinal alternatives—the possibilities over which attitude verbs quantify—not with possible worlds, but with centered worlds, where a centered world is a triple consisting of a possible world, a time, and an individual (the center), who exists at the time and world in question. A centered world $(w', t', x')$ will be compatible with what Lingens believes (at time $t$ in world $w$) if and only if $x'$ has, at $t'$ in $w'$, all the properties that Lingens believes $de$ $se$ that he himself has at $t$ in $w$. The content of a belief, on the centered worlds proposal, is a centered proposition, a set of centered worlds (or the characteristic function thereof). A centered proposition $p$ is true at a centered world $(w, t, x)$ iff $(w, t, x) \in p$ (or $p(w, t, x) = 1$). Lingens believes a centered proposition $p$ at $t$ in $w$ iff $p$ is true at all the centered worlds compatible with what Lingens believes at $t$ in $w$. If Lingens believes a centered proposition $p$ at $t$ in $w$, his belief is true simpliciter iff $p$ is true at $(w, t, Lingens)$. (2012, p. 5:8)

In its most extreme developments (as in Ninan 2012), centered-worlds talk becomes extended to multi-centered-worlds talk, in which contents are sets of worlds coupled with an entire sequence of directly-given objects:

For Lewis, the basic notion of belief is belief $de$ $se$. This is the basic notion in two (related) senses: First, this is the only notion of belief that one needs when assessing a potential doxastic alternative for compatibility with what an agent believes. Second, the notion of $de$ $re$ belief—believing something about someone relative to an acquaintance relation—is defined in terms of the notion of $de$ $se$ belief. On the multi-centered approach, in contrast, the basic notion of
belief is belief *de re*, or the notion of what an agent believes about an individual relative to an acquaintance relation, or, more generally, what an agent believes about a plurality of individuals-relative-to-acquaintance-relations. Again, this is the basic notion in two senses. First, it is this notion of belief that we use when we characterize the notion of doxastic compatibility. Second, *de se* belief is understood as a special case of *de re* belief—it is a *de re* belief where the res is the agent and the relation of acquaintance is identity. (2012, pp. 5:37–38)

It is hard, in reading Ninan’s development of the Lewisian picture, to avoid thinking that we have come full circle: we started with a fundamentally descriptivist/ *de dicto* picture, added a single referential point in an attempt to preserve the broadly descriptivist character, and worked our way back to a traditional referentialist picture on which the standard cognitive orientation to the world is *de re* /object-dependent, and in which indexicality no longer plays any distinctive role.

5.4. Response 3 to the Two Gods: It’s Still All Frege Puzzles

Recall that for Lewis’s two gods argument to work, it is crucial that the propositionally omniscient gods still not know who and where they are. Suppose Zeus is the god on the tallest mountain and Odin is the god on the coldest mountain. Then Zeus, despite his propositional omniscience, does not know that he is the god on the tallest mountain. He does, however, presumably know that *Zeus* is the god on the tallest mountain. There is a straightforward possible-worlds content to be assigned to the belief that Zeus is the god on the tallest mountain. It is not Lewis’s preferred content for that belief, after the switch to the *de se* framework, but it is available, and if Lewis’s argument for the *de se* framework is not to be question-begging, it needs to allow that the gods believe that content, and thereby believe that Zeus is the god on the tallest mountain.

So what Lewis is trying to explain is how Zeus can know that *Zeus* is the god on the tallest mountain, but not know that *he* is the god on the tallest mountain. But this, of course, is just a special case of traditional Frege puzzles. As we have mentioned in Chapter 4, the
problem that Lewis is concerned with is a general one of understanding how contents can be more fine-grained than sets of worlds and how propositional attitudes can be sensitive to such fine-grained distinctions in content. If that is indeed the general problem, it is surely an adequacy condition on an account on a theory of opacity that it be fully general. We don’t want to have one theory explaining how one can believe Superman flies without believing Clark Kent flies, and a different one explaining how one can believe Hesperus is a planet without believing Phosphorus is a planet. Similarly, we don’t want one theory explaining how one can believe Zeus is the god on the tallest mountain without believing Jupiter is the god on the tallest mountain, and a different one explaining how Zeus can believe Zeus is the god on the tallest mountain without believing he is the god on the tallest mountain.

Of course, Lewis does have a general account of opacity, but it is a general account that gives indexicality and the de se a special and central place. So the question is: why think that problems of opacity call for this kind of indexicalized content? Lewis does not give us any reason to think that his indexical versions of Frege puzzles are special in any way, so we’re left with no reason for thinking that the solution to Frege puzzles is to say that we think of all objects indirectly, by thinking about ourselves.

Again, that’s not to say that there is a different solution to Frege puzzles that we prefer—it’s just to say that we’ve been given no reason to think that a special treatment of indexicality is the right solution. Maybe indexical versions of Frege puzzles do have to be treated differently from other versions; maybe the right thing to say in the end is that the solution to Frege puzzles is a strange cognitive egomania. But no argument that Lewis or others give shows that that is the case.

5.5. Response 4 to the Two Gods: Worlds Are Not All There Is

As emphasized earlier, for Lewis’s two gods argument to be run without begging the question, it must be coupled with a theory of content. Moreover, that theory of content must make good on the thought
that one can know all of those contents without knowing that one is on the tallest mountain. If the theory of content is too fine-grained, it will fail Lewis’s dialectical needs—if there are distinct propositions *that Zeus is the god on the tallest mountain* and *that I am the god on the tallest mountain* (as thought by Zeus), then Zeus, in knowing all propositions, will know that he is the god on the tallest mountain.

So it’s no accident that Lewis starts with a possible-worlds framework. And not just any version of that framework—it’s important that it be specifically *metaphysically* possible worlds. Were there worlds in which Zeus was the god on the tallest mountain, but he (demonstrating Zeus) was not the god on the tallest mountain, then we would have a more fine-grained set of possible-worlds contents, and again Lewis’s problem could not get off the ground. For the problem to get started, we need a notion of worlds which are closed under co-reference, so that if \( a \) is \( b \) in \( w \), and \( a \) is \( F \) in \( w \), then \( b \) is \( F \) in \( w \). Lewis’s modal realism, of course, provides one such picture of worlds.

Possible-worlds accounts of contents are notoriously coarse-grained. They fail to differentiate between logically or mathematically equivalent contents (*that* \( 3 + 7 = 10 \) is the same content as *that* \( 123 + 456 = 579 \)). They fail to distinguish between metaphysically equivalent contents (*that Kripke is a philosopher* is the same content as *that Kripke is a human who is a philosopher*). They fail to distinguish between coreferential expressions (*that Superman flies* is the same content as *that Clark Kent flies*). And, as Lewis points out, they fail to distinguish between indexical and non-indexical presentations (*that Zeus is the god on the tallest mountain* is the same content as *that he is the god on the tallest mountain*). Advocates of possible-worlds accounts have things to say in response to each of these worries, of course, and it’s no part of our task to argue that these things fail. But what we do want to point out is that Lewis has a very narrow bit of logical space on which to rest his argument. He needs possible-worlds accounts to be viable enough that the choice of them as the tool for cashing out *propositions* is theoretically interesting. But he needs them to be *problematic* enough that they fail to make the distinction that he wants made. Again, we don’t argue that there is *no* such bit of logical space. But once one sees how narrow it is, it may be tempting simply to look for other theories of content.
Lewis is explicit about the dependence of his argument on the possible-worlds conception of content:

My target in this paper is the view, until recently my own, that the objects of attitudes are propositions in the sense of sets of worlds. **I need not quarrel with the view that they are propositions in some other sense.** (1979, p. 515, bolded emphasis added)

And he says of Perry’s view, on which in essence contents are *structured* propositions, that “I am sure it works as well as mine, but it is more complicated” (p. 537).

Indeed, it is easy to see that structured propositions must do at least as well as Lewis’s account in solving the informational puzzle of the two gods. Thinking of Lewis’s picture in centered-worlds terms, what Lewis gives us are contents that pair worlds with individuals. The content that Zeus does not grasp, and which will resolve his informational uncertainty about his location, is the content that pairs worlds with individuals living on the tallest mountains in those worlds:

\[ C_1 = \{ \langle x, y \rangle : x \text{ lives on the tallest mountain in } y \} \]

Structured propositions, of course, do not directly give us either worlds or individuals. But if structured propositions have truth conditions, they determine sets of worlds—namely, the worlds in which those truth conditions are satisfied. And because structured propositions are structured, they can contain individuals as constituents, and determine privileged individuals in that way.

So consider the structured proposition *that Zeus is the god on the tallest mountain*. This structured proposition determines a set of worlds: the worlds in which Zeus is the god on the tallest mountain. And by virtue of containing Zeus as a constituent, it determines a privileged individual: Zeus. So we can naturally associate with the structured proposition a set of centered worlds:

\[ C_2 = \{ \langle x, y \rangle : x = \text{Zeus and } y \text{ is a world in which Zeus is the god on the tallest mountain in } y \} \]

But \( C_2 \) is a proper subset of \( C_1 \). Given \( \langle x, y \rangle \) such that \( x \) is Zeus and Zeus is the god on the tallest mountain in \( y \), we have \( \langle x, y \rangle \) such that \( x \) lives on the tallest mountain in \( y \). So \( C_2 \) is strictly more informative.
than C1. If Zeus’ locative anxiety can be resolved by coming to know C1, it can a fortiori be resolved by coming to know C2. If centered worlds solve Lewis’s two gods problems, structured propositions must solve it as well, but without special indexical contents.

5.6. The Role of Self-Ascription

The case of the two gods, then, gives us no reason to accept a revisionary theory of content—it is, at best, just another instance of Frege’s puzzle, will be solved by whatever our general theory of opacity is, and points to no special indexical problem of opacity. What we argue now is that even if one accepts that there is a special problem to be solved here, it is not the revisionary theory of content that does the solving, but the unexplained notion of self-ascription of a property.

To say that contents are properties, rather than propositions, is so far to remove the notion of truth conditions from the theory of content, since properties, unlike propositions, are not the sort of thing that have truth conditions or are true or false. Lewis restores truth conditions by coupling properties-as-contents with believing-a-content as self-ascribing that property. The content I am the god on the tallest mountain is the property being the god on the tallest mountain; what it is to believe that content is to self-ascribe that property.

But this picture solves the problem of the two gods only because it is specifically self-ascription. Zeus, being propositionally omniscient, already ascribes to Zeus the property of being on the tallest mountain. What he needs to do to resolve his locative uncertainty is to ascribe to himself the property of being on the tallest mountain. So to understand the difference between knowing that Zeus is on the tallest mountain and knowing that he is on the tallest mountain, we need to understand the difference between (for Zeus) Zeus-ascribing and self-ascribing a property.

But that difference is entirely extrinsic to the theory of content Lewis is proposing. To see this, consider two other applications to which the Lewisian revisionary account of content could be put. Consider first Lewis*. He has the following puzzle:

Two gods are in conversation with one another. One sits on the tallest mountain and throws down manna; the other on the coldest mountain and throws
down thunderbolts. Each knows all propositional information about the world. But the god on the tallest mountain does not know to whom he is talking—is it the manna-thrower, or the thunderbolt-hurler?

Lewis* proposes the following solution:

Sentences take properties as contents. When I endorse the content of a sentence \( S \), I audience-attribute the property that its semantic content represents. The god on the tallest mountain needs to learn whether he should audience-attribute the property of manna-throwing or audience-attribute the property of thunderbolt-hurling.

It should be obvious that both Lewis and Lewis* use exactly the same formal machinery. Their collections of possible semantic contents are identical. The difference between their accounts lies not in the semantic values, but in the doxastic difference between self-ascription and audience-ascription. 7

Next consider Lewis**. He has the following puzzle:

The god on the tallest mountain is trying to figure out who Zeus is. He has a god’s-eye total descriptive view of the world, but that view doesn’t tell him who Zeus is.

Lewis** proposes to solve his puzzle by claiming that sentences carry properties as semantic values, in such a way that when one endorses a sentence \( S \), one thereby Zeus-attributes the property semantically assigned to \( S \). Since Zeus is the god on the tallest mountain, what Zeus needs to do is Zeus-attribute the property of being on the tallest mountain. 8

So self-ascription is not the only theory of ascription that could be coupled with a theory of properties-as-contents. Any theory of

---

7 That properties-as-contents could be coupled with a theory of audience-ascription to produce a theory of de tu attitudes does not, of course, entail that they should be. Perhaps this account is less general than Lewis’s, since it is applicable only to cases in which one is in conversation with someone. Perhaps it is less convenient, because we track ourselves more easily than we track others. We are not recommending this account; we are merely using it to highlight the separate roles of the theory of content and the theory of property ascription.

8 Again, we’re not endorsing this picture. Again, perhaps it is less general and convenient than Lewis’s full picture. But it highlights the difference between Zeus’ self-ascribing the property of being on the tallest mountain (we can imagine that Zeus knows himself to be on that mountain) and Zeus’ Zeus-ascribing that property.
ascription can be so used, but different puzzles are addressed through different theories. When we couple property contents with a theory of audience-ascription, we get a theory designed to address *de tu* attitudes. When we couple property contents with a theory of Zeus-ascription, we get a theory designed to address *de re* attitudes about Zeus. When we couple property contents with a theory of self-ascription, we get a theory designed to address *de se* attitudes. So a first lesson: it is only when combined with a theory of self-ascription that property/centered-worlds contents solve the problem of the two gods.

But we can say more than this. In fact, it is self-ascription that is doing *all* of the work. The switch from propositions to properties as semantic values is completely inert. To see this, consider the way in which a hypothetical Lewis*** solves the problem of the two gods. He begins by noting that the proposition that *Zeus is on the tallest mountain* is the same proposition as the proposition that *he is on the tallest mountain*. He then holds that what Zeus needs to do to solve his locative uncertainty is to *self-believe*, rather than *other-believe*, that proposition. Note the formal similarity with Lewis’s solution: in both cases, we have a single content, which is itself too coarse-grained to capture the distinctions we want. We then couple that content with a specialized cognitive relation (*self-ascribing* for Lewis, *self-believing* for Lewis*** ) that does the fine-graining and captures the “*de se*” nature of the attitude. If one approach solves the problem, the other will as well. But if the problem can be solved in the possible-worlds framework as easily as in the properties framework, then it is the self-ascription/self-believing that is doing the work.

We take it that many whose work is inspired by Lewis disagree with us about this, and hold that the Lewisian contents are doing important work. Here, for example, is a remark by Andy Egan, typical of the attitude we have in mind. Talking about Perry’s classic cases from “The Essential Indexical,” Egan says this about the difference between Lewis and Perry:

Perry draws an importantly different conclusion from the cases. He wants to locate the extra structure not in the contents of belief, but in the belief-states. So, when each of us believes that our own pants are on fire, there’s no common *content* to our beliefs, but we’re in the same kind of belief-state. If this kind of account is right, then we don’t need centered-worlds contents for
propositional attitudes. I’m going to assume without argument that this kind of account isn’t right, and that propositional attitudes do, at least sometimes, have centered-worlds contents. (And that since centered-worlds contents are the contents of propositional attitudes, it’s safe to call them propositions.) It’s worth noticing that if I’m wrong about the need for centered-worlds content, pretty much everything I say from now on will be false. I feel no pangs of conscience as I move on without further discussion because: (a) I’m pretty convinced that, at the end of the day, we will want to locate the extra structure at the level of content. . . . (2006b, p. 107)

Egan is here arguing that Lewis’s centered-worlds framework can help provide an account of the distinction between primary and secondary qualities, but the point we want to draw attention to is the claim that in a centered-worlds framework, the “extra structure” is supposedly located in the content. That simply isn’t Lewis’s view. Self-location was what did all the work for Lewis.

Our modified “possible worlds” Lewisian made crucial use of the notion of self-believing to solve the problem of the two gods. One might worry, however, that we haven’t told you what it is to self-believe a proposition (or, what the difference is between self-believing and other-believing a proposition). That’s perhaps a fair point, but it looks to be as applicable to Lewis as it is to us. Despite the centrality of the notion of self-ascription to Lewis’s account, Lewis has very little to say about what it is to self-ascribe a property. We take it that there are two plausible Lewisian moves at this point.

First, Lewis could take self-ascription to be elucidated through examples. The thought here is that we know what the difference is between Zeus’ knowing that Zeus is on the tallest mountain and Zeus’ knowing that he is on the tallest mountain, and that we reverse-engineer a conception of self-ascription from our grasp of that difference. Self-ascription is that relation Zeus stands in to the property of being on the tallest mountain when he believes himself to be on the tallest mountain.

Two points about this route. First, self-belief can clearly be elucidated just as effectively as self-ascription using this method, so going this way only highlights the conclusion that it isn’t the theory of content that’s doing substantive work for us. Second, to go this route is to return the discussion to the ground covered in Chapters 3–4. In those
chapters, we have argued that there is no distinctive difference between Zeus’ knowing that Zeus is on the tallest mountain and Zeus’ knowing that he is on the tallest mountain—that if there is a difference, it is just another instance of Frege’s puzzle.

The second option for Lewis is to rely on a general functionalist theory of mind to supply the account of self-ascription. Thus, for example, Seth Yalcin says: “talk of ‘self-ascription’ within Lewis’s framework is a dispensable heuristic, not something carrying an explanatory load, and not something essential to understanding the proposal” (2012, p. 3). Talk of self-ascription, according to Yalcin, is a way of talking about functional role:

The foundational story for Lewis’s model of de se content builds on the idea that we are constitutively rational: the intentional mental states of an agent have the content that they have largely in virtue of the fact that the dispositions of the agent to act can be explained as rational on the hypothesis that the agent’s mental states have that content.

[To believe B and desire D, where B and D are sets of centered worlds, is (at least) to be disposed to perform the sorts of actions which would, if performed by an agent located in B, tend to bring about it about that the agent is in D]. 9

Crude, obviously, but it does tells us a great deal about what it is supposed to mean for a centered world to be compatible with one’s beliefs—what it means, on Lewis’s model, to self-ascribe a property. It is (largely) to be in a state of mind with this sort of functional structure. (2012, p. 3)

We take this objection very seriously and so will devote the entire next chapter to addressing it. We will, however, address this not just as a point about Lewis exegesis, but as a more general objection to our deflationary account of the de se and indexicality.

5.7. Brief Summary of This Chapter

Here are four take-home messages from this chapter:

1) Lewis presents another kind of argument in favor of Essential Indexicality. He gives an informational puzzle, asking what sort of

9 We here interpolate the corrected version of the functionalist specification from fn. 2 of Yalcin’s handout into the main text.
information could be gained by one of the two gods when that
god learns his location. He then proposes a specific kind of per-
spectival de se content to solve this puzzle—mental state contents
are properties/centered worlds, rather than propositions. This content
is perspectival because it delivers truth conditions not absolutely,
but only relative to a choice of agent/center.

2) Lewis’s presentation of centered-worlds contents comes with a
compelling rhetoric about information as locating yourself in
local as well as modal space. But that rhetoric isn’t a serious argu-
ment—it relies, at bottom, on an optional decision to cast the
issues in terms of truth at a contextually given index rather than in
terms of truth simpliciter.

3) Disentangled from the rhetoric, Lewis’s informational puzzle is
just another instance of a Frege puzzle, and no reason is given to
think that this particular instance of opacity is importantly differ-
ent from other instances.

4) Lewis’s own solution to his informational puzzle proceeds by
combining his novel theory of properties-as-contents with a
specific view of the cognitive role of those contents—that to
believe a property-content is to self-ascribe the content. The role
of self-ascription is little emphasized either by Lewis or in the
subsequent literature. But the work of solving the informational
puzzle rests entirely on the self-ascription leg of the proposal, as
can be seen by the fact that equivalent solutions can easily be given
by combining other theories of content (such as possible-worlds
content) with analogous attitudes such as self-believing.
Chapter 6

Functionalism to the Rescue?

Our central theme in this book has been that the vaguely delimited class of phenomena that people lump into the category of the “de se” gives us no compelling reason to think that there is a philosophically deep role essentially played by indexical contents. Some will worry though that our focus so far has been too narrow. Perhaps we need to think about the philosophical foundations of the theory of content in order to see how the theory of content plays the proper “de se” role. In particular, it can appear tempting to appeal to an underlying functionalism to extract a revisionary theory of content from the “de se” phenomena. Call this the rescue-by-functionalism strategy. We have seen versions of this strategy make brief appearances already in both Chapters 3 and 5. A natural reaction to our arguments against indexicality-agency connections in Chapter 3 is to say that if the theoretical role of belief and desire contents is to explain action, then we can start with a list of cases in which action results, and then build a theory of mental content off of those cases, with “de se” contents being reverse-engineered as the contents that produce action. In Chapter 5, we saw this thought: Lewis’s picture of properties-as-contents needs to be coupled with an account of self-ascription of a content in order to produce a theory of the “de se,” but Lewis doesn’t provide one, at least not explicitly. So our understanding of what it is to self-ascribe a property will fall out of a general project of understanding mental contents via their functional role in folk psychology.
6.1. Attack vs. Defense

That’s a very sketchy and high-level picture of a course of action, but the devil is in the details. We begin by distinguishing two versions of the rescue-by-functionalism strategy:

**Attack.** A functionalist metasemantics will provide a positive argument in favor of a revisionary indexical or *de se* theory of content.

**Defense.** A functionalist metasemantics will not provide a positive argument in favor of a revisionary theory of content, but will serve to complete the revisionary account, showing how it is defended against various objections.

In “Reduction of Mind,” Lewis endorses Attack:

What determines the content of belief and desire?—The occupation of folk-psychological roles by physical states, presumably neural states; and ultimately the pattern of coinstantiation of fundamental physical properties and relations. (1999, p. 308)

My own approach can be summed up quickly. The contentful unity is the entire system of beliefs and desires.... That system is an inner state that typically causes behavior, and changes under the impact of perception (and also spontaneously). Its content is defined, insofar as it is defined at all, by constitutive rationality on the basis of its typical causal role. This content is in the first instance narrow and *de se*. (1999, p. 324)

Despite Lewis’s endorsement of the approach, we will first show that Attack (in the sense required to engage with our project) is not just incorrect, but completely and obviously hopeless. We note that rebutting Defense is not necessary for our purposes. Our goal is to show the *inessentiality* of indexicality by showing that no revisionary indexicalized theory of content is necessary. It does not follow from that position that there is no coherent indexicalized theory of content, and in particular that some such theory of content cannot be internally defended from objections. It follows only that we are not obligated to adopt such a theory, even if there is one. Nevertheless, we also argue that Defense is unsuccessful.
6.2. Problems for Attack

To see what is wrong with Attack, it is helpful to distinguish two versions of functionalism. We start with Folk Functionalism. According to Folk Functionalism, mental states are theoretically defined by Ramsifying on our folk-psychological theory. We cannot, of course, know exactly what the resulting definitions will look like without having on hand an actual specification of folk-psychological theory. It is worth keeping in mind when considering functionalist proposals that no one has ever produced such a specification and that it is far from clear that there is any unique thing, or perhaps even any very well delimited class of things, that counts as a “folk-psychological theory.” Moreover, in the absence of a specification of a folk-psychological theory, we should be wary of the claim that the Ramsified clauses will manage either to individuate roles so that, for example, the satisfiers of the “desire” role are consistently distinct from the satisfiers of the “belief” role, or to specify states that support the truth of ordinary claims about mental attitudes.

But we can give at least a dummy version of a Folk Functionalist definition, and the central problem can be observed even with the dummy version. So suppose we have:

\[(FoFu)\] An agent $A$ believes that $p$ and desires that $q$ if and only if, whenever there is an action $C$ such that performing $C$ when it is the case that $p$ tends to promote $q$, $A$ will ceteris paribus perform $C$.

\[1\] Because the goal of (FoFu) is to characterize mental states using the tools of folk psychology, specific instances of (FoFu) should use the language of folk psychology. So we should end up with instances like:

(1) Jones believes that the water is in the glass and desires that he drink water if and only if, whenever there is an action $C$ such that performing $C$ when it is the case that the water is in the glass tends to promote that he drink water, Jones will ceteris paribus perform $C$.

rather than:

(2) Jones stands in the belief relation to the proposition that the water is in the glass and in the desire relation to the proposition that he drink water if and only if, whenever there is an action $C$ such that performing $C$ when it is the case that the proposition that the water is in the glass is true tends to promote the truth of the proposition that he drink water, Jones will ceteris paribus perform $C$.

(FoFu) is thus best thought of as a schema in which $p$ and $q$ are replaced with particular sentences, yielding particular belief attributions. To write (FoFu) in the style of (2), rather than
We simplify in many ways here. (FoFu) specifies the content of a par-
ticular belief and desire, rather than of the entire belief-desire profile
of an agent, as many functionalists would prefer. It also makes no men-
tion of various complicating hedges on the dependence between the
satisfiers of the Ramsification of folk psychology and the mental states,
such as appeals to naturalness or constitutive rationality. None of these
simplifications affect the points we make.

Here’s the problem: nothing helpful about the shape of a theory of
content follows directly from such a definition. The language of con-
tent has not been used anywhere in the definition. (FoFu) says noth-
ing about possible worlds, structured propositions, centered worlds,
Fregean senses, interpreted logical forms, or any of the hundreds of
other formal devices that have been posited as vehicles of content.
Thus it does nothing directly to mandate one rather than another of
these, and as we will see, also too little indirectly to make interesting
choices between theories of content.

This should come as no surprise since the primary goal of the Folk
Functionalist account is metaphysical: it is to help us to figure out what
kinds things mental states are—they are those states that play a certain
role as specified by folk psychology. Pinning down the metaphysics
of mental states isn’t done by talking about content, and the ordinary
talk being used to pin it down is not talk couched in the theoretically
loaded vocabulary of the theory of content. Folk Functionalism gives
us a tool to determine that an agent $A$ believes that $p$, desires that $q$, and
so on (by considering the behavioral dispositions of $A$). If it succeeds,
it tells us what it is to so believe and desire—it is to be so disposed, or
to have realizing structures for those dispositions. But what it does
not tell us is what it is for a theoretical entity to be the content of $A$’s
belief or desire.

Consider two ways in which one might hope to extract a theory of
content, or at least some interesting constraints on a theory of content,
from Folk Functionalism. First, one might hope that some constraints
on the fineness of grain of contents could be imposed by the interac-
tion between Folk Functionalism and the theory of content. Perhaps

(1) is to shift to what we later call Fancy Functionalism. This fact is relevant to the discussion
of ternary theorists—see fn. 2. Thanks to Andy Egan for pressing us on this issue.
we can rule out possible-worlds contents, because possible-worlds contents would assign the same content to mental states that have distinct functional roles. If the platitudes of folk psychology lead us to attribute different *ceteris paribus* action profiles to agents who believe that $2 + 2 = 4$ and agents who believe that the negation of the Generalized Continuum Hypothesis is consistent with ZFC, then it looks like (FoFu) will require that different contents be assigned to those two beliefs, contrary to the dictates of a possible-worlds account of content. So by this line of thought, what Folk Functionalism tells us is that our formal account of content must be at least as fine-grained as the attributions of attitudes in folk-psychological explanations. At most what we are going to learn from functionalism, on this way of thinking about things, is that contents are very fine-grained. But, as we have emphasized in Chapter 4, this is a *generic Fregean point* about contents— *lots* of substitutions fail to preserve content, on this way of thinking about contents, and there’s no reason to think that *indexical* contents in particular are special.

But even this is too quick, as it presupposes an optional story about the relation between contents and mental-state individuation. For example, anyone who endorses a “ternary” account of belief, on which belief is a three-place relation among an agent, a belief, and a mode of presentation, will reject that story. A ternary theorist has no reason to follow all the way to the end the path that moves from (a) fineness of grain of functional role to (b) fineness of grain of mental states, and then from there to (c) fineness of grain of content. More generally: the word “content,” or theoretical analogues thereof, appears nowhere in (FoFu). Any move from functionalism to the theory of content—even a minimal move that just extracts some fineness of grain constraints on contents from the functionalism—requires auxiliary hypotheses.

---

2 Note that the ternary theorist agrees that (returning to the example of fn. 1) Jones believes there is glass in the water, and desires that he drink water, and that these two facts explain his action. The ternary theorist can thus accept (FoFu). He just doesn’t accept that, e.g., believing that there is glass in the water is the same thing as standing in the belief relation to the proposition that there is glass in the water, and hence cannot directly extract consequences for the theory of content (here, the individuation conditions for Russellian propositions) from the dictates of (FoFu). Were (FoFu) written in the style of (2) in fn. 1, with explicit mention of standing in the belief relation to a proposition, he could extract such conclusions. The ternary theorist presumably would not then endorse the modified (FoFu).
linking the formal notion of content to the ordinary practice of mental-state ascriptions.

A second way one might hope to extract constraints on a theory of content from Folk Functionalism: one might think that our folk psychology already suffices to pick out a privileged class of beliefs, which will yield a privileged class of contents. Perhaps there are certain beliefs that play a special role in the explanation of action, and hence Folk Functionalism will assign to them a special kind of content. The previous cautionary comments about the relation between mental states, as characterized by (FoFu), and the theory of content, continue to apply. But another point is crucial here. Maybe there is a class of beliefs that plays a special role in the explanation of action. Chapter 3 was largely devoted to arguing against the existence of such a class (or at least any interesting such class), and if Chapter 3 is successful, then this functionalist gambit fails. But maybe Chapter 3 fails, and there is such a class. However, it is a large additional step from this point to a specifically indexical theory of content. We need some special reason to think that the special category of beliefs will be related in some interesting way to the phenomenon of indexicality. In particular, the strategy of simply equating the action-motivating mental states with the first-person, or otherwise indexical, mental states will not succeed, since Folk Functionalism doesn’t allow this kind of direct equation—it’s just an open empirical question about our folk psychology whether it equates these two classes.

So, in brief, according to Folk Functionalism, the functional definitions are wholly without the technical vocabulary of any theory of content, revisionary or otherwise. No theory of content, revisionary or otherwise, will then be forthcoming from the functionalism. If Attack is to have any hope, it must be by way of Fancy Functionalism. Fancy Functionalism couches the functionalist definitions in a theoretically loaded vocabulary, already introducing the talk of content into the definitions. So, for example:

(FaFu) An agent A bears the belief relation to a set p of centered worlds and the desire relation to a set q of centered worlds if and only if, whenever there is an action C that would increase the likelihood of the ordered pair <A, actual world> being an
element of $Q$ on the condition that the ordered pair $<A, \text{actual world}>$ is an element of $P$, $A$ will *ceteris paribus* perform $C$.$^{3,4}$

Stalnaker, for example, is a Fancy Functionalist when he gives the functional specification of belief and desire in terms of the theoretic apparatus of possible worlds:

To desire that $P$ is to be disposed to act in ways that would tend to bring it about that $P$ in a world in which one’s beliefs, whatever they are, were true. To believe that $P$ is to be disposed to act in ways that would tend to satisfy one’s desires, whatever they are, in a world in which $P$ (together with one’s other beliefs) were true.$^5$ (1984, p. 15)

---

$^3$ Some commenters have suggested to us that no one understands functionalism in the Fancy Functionalist way. If you agree, feel free to skip ahead to the discussion of Defense, and keep in mind that the prospects for the functionalist strategy depend entirely on the prospects of Folk Functionalism. We include the discussion of Fancy Functionalism because we think it presents an option that ought to be attractive to some functionalists.

$^4$ Andy Egan has pointed out to us that this formulation of (FaFu) has various problematic features—it at least prima facie relies on the problematic idea of increasing the likelihood of the apparent necessary truth or falsehood that a certain ordered pair is in a certain set, and it has difficulties in accounting for cases in which $A$’s beliefs are incompatible with necessary properties of $A$. While we agree that these are legitimate concerns, none of what we will go on to say about (FaFu) depends on these details, so we omit the complications needed to avoid the worries.

$^5$ Stalnaker’s Fancy Functionalism then plays a crucial role in his extraction of a *coarse-grained* account of propositions from his ontology. He says:

It is clear from the general schemas for the definitions of those relations [belief and desire] that the following will be true: if the relation holds between an individual and a proposition $x$, and if $x$ is necessarily equivalent to proposition $y$, then the relation holds between the individual and $y$. This implies that the thesis that necessarily equivalent propositions are identical—the main substantive consequence of the possible worlds analysis of propositional content—is a thesis that is tied to, and motivated by, the causal-pragmatic explanation of intentionality. (1984, p. 24)

If the functionalism is cast in purely Folk Functionalist terms, the theoretical vocabulary of propositions and worlds will not enter into it, and we will not be able to extract from the coarse-grainedness of the functional roles anything about the grain of propositions. Note that Stalnaker’s Fancy Functionalism is relatively minimal—the coarse-graining of propositions requires only the move between *$A$ believes that $p$* and *$A$ stands in the belief relation to the proposition that $p$*. 
Fancy Functionalism at least has the right form for achieving Attack. If a definition like (FaFu) is the right definition of belief and desire, then it will indeed mandate a revisionary theory of content.⁶

6.3. Two Problems for Fancy Functionalism

But there are two important problems with Fancy Functionalism: the arbitrariness problem and the circularity problem. Consider first the arbitrariness problem. (FaFu) in Section 6.2 is written using a centered-worlds theory of content. But it looks trivial to rewrite versions of (FaFu) for other theories of content. Thus, consider also:

(FaFu’) An agent A bears the belief relation to a structured proposition p and a structured proposition q if and only if, whenever there is an action C that would increase the likelihood of the truth of q on the condition of the truth of p, A will ceteris paribus perform C.

(FaFu’’) An agent A bears the belief relation to a set of possible worlds p and a set of possible worlds q if and only if, whenever there is an action C that would increase the likelihood of the actual world being an element of Q on the condition that the actual world is an element of P, A will ceteris paribus perform C.

Indeed, the availability of reworkings along the lines of (FaFu’) and (FaFu’’) seems inevitable, given “analytic functionalism” in particular. According to analytic functionalism, the functionalist definitions are analytic truths. But surely the only way that a definition using theoretical jargon, be it centered worlds or structured propositions, is going to be analytic is if it is derived from an underlying pre-theoretical folk-psychological account via a translation procedure between

⁶ More or less: there remains, because of the possibility of ternary accounts of belief, a worry about the connection between being the object of the belief relation and being a content. Even more theoretically loaded versions of Fancy Functionalism could work around this worry, making them increasingly vulnerable to the objections listed in Sect. 6.3.
content-theoretic and pre-content-theoretic vocabulary. So the general form is to move from (FoFu) to a schematic (FaFu*):

(FoFu) An agent $A$ believes that $p$ and desires that $q$ if and only if, whenever there is an action $C$ such that performing $C$ when it is the case that $p$ tends to promote $q$, $A$ will *ceteris paribus* perform $C$.

(FaFu*) An agent $A$ stands in the belief relation to an entity $p$ and stands in the desire relation to an entity $q$ if and only if, whenever there is an action $C$ such that performing $C$ when things are with respect to $p$ the way they need to be for the belief of $A$’s that is constituted by standing in the belief relation to $p$ to be veridical tends to promote things being with $q$ the way they need to be for the desire of $A$’s that is constituted by standing in the desire relation to $q$ to be satisfied, $A$ will *ceteris paribus* perform $C$.

The schematic (FaFu*) is then realized in various forms—(FaFu), (FaFu’), (FaFu’’)—by filling in simultaneously the details of the nature of the entities and the way things need to be with respect to those entities. But there seems to be nothing deriving from the functionalism allowing us to choose from among the various fillings—in of the schematic (FaFu*). (Maybe you think there’s some *other* role that contents play that will allow the choice, but then the argument for, or explanation of, Essential Indexicality is coming from that somewhere else, not from the functionalism.7)

Of course, if we adopt a non-analytic functionalism, there might be an argumentative route to one version of (FaFu) over the others. But it’s far from obvious what that route would be, and in particular, we don’t think that anyone has presented an argument that their preferred theory of content integrates uniquely well into a functionalist metasemantics.

Also, the arbitrariness objection has bite twice over. First, as already noted, one theory of content can be swapped out for another. Fancy Functionalist definitions using centered worlds look easily reworked

---

7 See, for example, Motivation 3 in Ch. 3. One thing you might want your theory of content to do is capture certain generalizations about the explanation of action. But that goal is independent of the functionalist story about mental state contents, and so that argument for essential indexicality can be assessed independently of the issue of functionalism.
using structured propositions, possible worlds, or whatever your favorite theory of content may be. This is not to say that theories of content are always trivially interchangeable. Indeed, one of our themes throughout has been the superiority of one account of content (a traditional non-perspectival account) over another (an account with some form or other of de se contents). It is just to say that where there is superiority, it is not something that will simply fall out of a functionalist specification of mental states—in part because, again, the goal of such functionalist specification is to give the metaphysics of the mental in terms of its causal and explanatory integration with ordinary psychological attribution practices, rather than to give a theoretical account of the role of content. If we build an account of content into the psychological account specifying functional roles—as in non-analytic Fancy Functionalism—then we also lose any hope that the functionalism per se will play a useful argumentative role. The debate then becomes one over the proper choice of theory on which to Ramsify, and that debate will simply recapitulate the debates occupying the earlier chapters of this book.

But second, even given a choice of content-types, there will be more than one way to write Fancy Functionalist clauses using that type of content. In abstraction, this point is obvious. Thus compare:

(FaFu´´) An agent $A$ bears the belief relation to a set of possible worlds $p$ and a set of possible worlds $q$ if and only if, whenever there is an action $C$ that would increase the likelihood of the actual world being an element of $q$ on the condition that the actual world is an element of $p$, $A$ will ceteris paribus perform $C$.

(FaFu´´´) An agent $A$ bears the belief relation to a set of possible worlds $p$ and a set of possible worlds $q$ if and only if, whenever there is an action $C$ that would increase the likelihood of the actual world not being an element of $q$ on the condition that the actual world is not an element of $p$, $A$ will ceteris paribus perform $C$.

Just deciding to use possible-worlds contents in our functionalist clauses does not yet tell us what role those contents need to play. We can characterize an agent’s goal of inquiry as being to determine what
world he is in, or to determine which worlds he is not in. Both characterizations will integrate equally well with the background functionalism. (Of course, the relation between contents and the semantics of belief ascriptions will need to be modified depending on which of these two options we pick.)

That’s a boring version of the point, but there are more interesting versions. On the centered-worlds version of Fancy Functionalism, the goal of an agent A’s inquiry is characterized as determining which properties A himself instantiates. But, as we noted in the discussion of Lewis earlier, properties can be used differently in statements of the goals of inquiry. So we could have:

\[(FaFu')\] An agent A bears the belief relation to a set \(p\) of centered worlds and the desire relation to a set \(q\) of centered worlds if and only if, whenever there is an action \(C\) that would increase the likelihood of the ordered pair \(<\text{François}, \text{actual world}>\) being an element of \(q\) on the condition that the ordered pair \(<\text{François}, \text{actual world}>\) is an element of \(p\), A will \(ceteris paribus\) perform \(C\).

Here the agent’s goal in inquiry is being characterized as the determination of what properties François instantiates. This needs some massaging, since this cannot be A’s goal in inquiry in worlds in which François does not exist. But the massaging is not difficult (for example, it can be achieved by disjunctive specifications of the goal of inquiry). And if the agent’s goal is characterized in this way, there is—despite the use of centered worlds, and despite the anchoring of the centered-worlds talk in a functionalist metasemantics—no genuine \(de se\) or perspectival aspect of the theory of content.

So the arbitrariness objection has bite twice over. First, the choice of one theory of content rather than another to write into functionalist definitions appears arbitrary. Second, the choice of larger theoretical role for a given content to write into functionalist definitions also appears arbitrary. Given this, Attack is doomed to failure. There is no way that functionalism can anchor a positive argument for a \(de se\) conception of content if it is merely an arbitrary choice to write the functionalist clauses in the \(de se\) rather than a non-\(de se\) way.
6.4. Back to Defense

What remains, then, is Defense. The thought behind Defense is that the functionalist metasemantics will provide the final piece of internal conceptual closure for the de se story. So, for example, the Lewisian account of the de se makes use of centered worlds, and accounts for the (putatively) characteristic features of de se mental states by taking those mental states to carry centered-worlds content. But it’s not enough to say that the contents of mental states are centered worlds; it’s crucial that the mental states are characterized by the right relation to the centered-worlds content. So what’s crucial is that the agent A self-ascribe, rather than (for example) audience-ascribe, or François-ascribe, or (crucially) A-ascribe the appropriate set of centered worlds. The account of self-ascription is then to be derived from the functionalism. Belief, in particular, is the mental state that involves self-ascribing, rather than (in one form or other) other-ascribing centered-worlds states. So examination of the functionalist definitions will allow us to see the difference between self-ascription and other-ascription (and, again crucially, the difference for A between self-ascription and A-ascription).

Defense is, of course, immune to the arbitrariness objection. The philosophical aspirations of Defense are not to provide a positive reason to adopt one form of functionalist definition over another, but just to show that a given de se theory of content can be supported by an appropriate functionalism. But a second line of objection to Attack is equally potent against Defense. This is the circularity objection.

Consider first an analogous case. Fregeans posit a difference in content between co-referential names with different cognitive significance, such as “Hesperus” and “Phosphorus.” Can we use functionalism to extract an account of the difference in content between co-referential names? Suppose we take our schematic Fregean functionalist definition to be:

\[(FrF) \ A \text{ believes that } p \text{ and desires that } q \text{ if and only if, when an action } C \text{ would promote } q \text{ on the condition that } p, A \text{ will ceteris paribus perform } C.\]

To achieve a difference between Hesperus-contents and Phosphorus-contents, we need a difference between (FrF1) and (FrF2):
(FrF1) A believes that Hesperus is F and desires that q if and only if, when an action C would promote q on the condition that Hesperus is F, A will ceteris paribus perform C.

(FrF2) A believes that Phosphorus is F and desires that q if and only if, when an action would promote q on the condition that Phosphorus is F, A will ceteris paribus perform C.

But for there to be such a difference, there needs to be a difference between an action promoting q on the condition that Hesperus is F and promoting q on the condition that Phosphorus is F, which means there needs to be a difference between the condition of Hesperus’ being F and the condition of Phosphorus’ being F. But now there is a disjunctive worry.

**Case 1.** The occurrences of “Hesperus” and “Phosphorus” on the righthand side of the functionalist clauses (FrF1) and (FrF2) are in extensional contexts. In that case, Fregeans and non-Fregeans alike will agree that they have the same truth value, and hence the choice of proper functionalist analysis cannot be governed by mere considerations of truth.\(^8\) Instead, what the Fregean will need is that the contents of “Hesperus” and “Phosphorus” differ in (FrF1) and (FrF2), and that it is this variation in content that allows (FrF1) and (FrF2) to ground the Fregean theory of content.

But it is hard to see how any explanation or grounding has been achieved here. It amounts to nothing more than capturing the difference in content by explaining that “Hesperus” refers to Hesperus, while “Phosphorus” refers to Phosphorus. Some Fregeans seem to have felt that such claims can, in fact, illuminate the nature of sense,\(^9\) but it is mysterious to us what the nature of the illumination is.

**Case 2.** The occurrences of “Hesperus” and “Phosphorus” on the righthand side of the functionalist clauses (FrF1) and (FrF2) are in hyperintensional contexts. (Perhaps the definitions should instead read “when an action C is taken by A to promote q on the condition...”.) There is a worry about the non-circularity of the functionalist definitions when attitude contexts

---

\(^8\) Equivalently, the putative insight granted by the functionalism into the content-difference between the co-referential “Hesperus” and “Phosphorus” cannot be achieved via the truth rather than falsity of (FrF1) or (FrF2).

\(^9\) See, e.g. the discussion of Ludlow’s view in Ch. 4.
are used on the righthand side, but even setting aside this worry, it is again hard to see how there is explanation or grounding. It is true that there will be more widespread agreement that there is a difference between, e.g. taking Hesperus to be \( F \) and taking Phosphorus to be \( F \), but the success of the functionalist definition now depends on a prior grip on what that difference in attitude amounts to. But the aspiration of the Fregean was to explain the difference in attitude via a difference in the content of the attitude, or more humbly, to explain the difference in the content of the attitude by a difference in the attitude, and to explain the difference in the attitude by a difference in the functional role of the attitude. Neither aspiration, however, can be achieved here.

The same dialectic applies to the \( de \ se \) cases. The functionalism is intended to explain or ground the peculiar first-person, perspectival, or indexical nature of the contents or states, but in the end will only presuppose a grasp of that nature. Consider a representative case. Stanley also objects that Lewis’s account of the \( de \ se \) rests on an unexplained notion of self-ascription:

The self-ascription framework claims to explain \( [de \ se \ \text{phenomena}] \), by appeal to the fact that John wants to self-ascribe a property. But this is not an explanation. What we desired was an explanation of what it means to self-ascribe, not just using that vocabulary in the meta-language. (2011, p. 89)

Yalcin’s response is that “talk of ‘self-ascription’ within Lewis’s framework is a dispensable heuristic”—dispensable in favor of a functionalist metasemantics. Yalcin thus offers:

To believe \( B \) and desire \( D \), where \( B \) and \( D \) are sets of centered worlds, is (at least) to be disposed to perform the sorts of actions which would, if performed by an agent located in \( B \), tend to bring about it about that the agent is in \( D \). (2012, p. 3, fn. 2)

Yalcin says, “Crude, obviously, but it does tell us a great deal about what it is supposed to mean for a centered world to be compatible with one’s beliefs—what it means, on Lewis’s model, to self-ascribe a property. It is (largely) to be in a state of mind with this sort of functional structure” (2012, p. 3). But the danger is that, in fact, it tells us nothing at all helpful about what it means to self-ascribe.
Making explicit the quantificational structure of Yalcin’s Fancy Functionalism, we have:

For any agent $x$ and any sets of centered worlds $B$ and $D$, $x$ stands in the belief relation to $B$ and the desire relation to $D$ iff: for any agent $y$, any action $A$, and any world $w$, if: if $y$ performs $A$ and $<y,w>$ is in $B$, then $<y,w>$ is in $D$, then $x$ performs $A$.  

Here is the heart of the explanatory failure: the functionalism makes repeated mention, under quantification, of an agent $y$. But it does not impose any constrain on the guise under which $y$ is presented, and without a requirement that $y$ is presented “in the de se way,” the functionalist clause cannot underwrite a story specifically of self-ascription, rather than of $x$-ascription by $x$.

We now consider how to make good on that general idea in the specific case of Yalcin’s functionalist clause. Following standard Perry-case reasoning, suppose François is identity-confused. He thinks about François twice over—for convenience, we’ll use the names $F_1$ and $F_2$ to represent François’s two ways of thinking about himself. We can suppose that $F_1$ is the “de se” way for François to think about himself, on the assumption that there is such a way. Consider the following ways in which François’s dispositions to act could be structured:

(G1) François undertakes actions $A$ when (all other things being equal) the fact that $F_1$ performs $A$ and $<F_2,w>$ in $B$ entails that $<F_2,w>$ in $D$.

(G2) François undertakes actions $A$ when (all other things being equal) the fact that $F_1$ performs $A$ and $<F_2,w>$ in $B$ entails that $<F_1,w>$ in $D$.

(G3) François undertakes actions $A$ when (all other things being equal) the fact that $F_1$ performs $A$ and $<F_1,w>$ in $B$ entails that $<F_2,w>$ in $D$.

10 This runs roughshod over a few of the subtleties of Yalcin’s clause. The “at least” proviso has been wholly dropped, actions that tend to bring about $D$ have been replaced with actions that simply do bring about $D$, and a disposition to perform a sort of action has been replaced with a universal performance of qualifying actions. None of these simplifications, however, matters to the point we will make.
(G4) François undertakes actions A when (all other things being equal) the fact that F1 performs A and <F1,w> in B entails that <F1,w> in D.

(G1)–(G4) are all ways that François can (partially) fulfill the condition of being an x who undertakes an action A when for any y and w, y’s undertaking A when <y,w> in B leads to <y,w> in D (because F1 and F2 are, in fact, the same individual). But they (putatively) differ in the way in which a de se requirement is placed on François’s conception of himself in his goal. Suppose François reads a diary of François (failing, of course, to recognize it as his own). Inscribed in the diary is the goal of becoming a famous philosopher. If he undertakes, guided by the characterization of the situation of he himself (i.e. François-as-F1), to make François (i.e. François-as-F2) a famous philosopher, his action-dispositions are characterized by (G3). If he undertakes, as guided by the characterization of the situation of François, to make François a famous philosopher, his action-dispositions are characterized by (G1). If, persuaded by the diary, he adopts as his own the project of becoming a famous philosopher, and pursues that project guided by the characterization of the situation of Francois, his action-dispositions are characterized by (G2). If, persuaded by the diary, he adopts as his own the project of becoming a famous philosopher, and pursues that project guided by the characterization of the situation of he himself, his action-dispositions are as characterized by (G4). Presumably only (G4) properly underwrites a notion of self-ascription—it is only in taking on the project to become a famous philosopher in the (G4)-way that one comes to desire to become a famous philosopher. So if the functionalism is to explain Lewis’s use of the language of self-ascription, we need to understand the difference between (G1)–(G4). We ourselves, as we have repeatedly stressed, are skeptical that there is a real and stable distinction to be made among (G1)–(G4), and even more skeptical that if there is a distinction, it is one that is interestingly and importantly local to indexical phenomena. But even setting these kinds of skepticism aside, the difference among these clauses is precisely the (putative) difference between the de se and the de re, so no progress has been made in explaining the de se. That’s not to say that the coupling of a functionalist metasemantics with a revisionary indexical theory of content...
has been rebutted—it’s just to say that the functionalism hasn’t added anything to our understanding of the distinctively indexical phenomenon, and that’s enough for the failure of Defense.

6.5. Brief Summary of This Chapter

Here are five take-home messages from this chapter:

1) It’s tempting to think that a functionalist account of mental state content will somehow undergird a project of setting out a special category of *de se* contents, by showing us something about what these contents are like, or how we are mentally related to them. In particular, Lewis’s notion of *self-ascription of a property* is intended by Lewis to be cashed out via functionalism.

2) If the functionalism is what we call Folk Functionalism, nothing helpful follows. The functionalism will deliver, for a given agent, a list of what is and is not believed. If we’re wrong in Chapter 4, and action essentially requires an indexical belief, then functionalism will ascribe indexical beliefs to each acting agent. But it won’t *explain* the *de se* nature of those beliefs—it will merely recapitulate the data. (And if we’re right in Chapter 3, it won’t even do that.)

3) If the functionalism is what we call Fancy Functionalism, then the theoretical results extracted from the functionalism are entirely a result of the theoretical commitments written into the specific form of Fancy Functionalism. If the functionalist clauses are written structured-proposition-style, structured propositions will come out. If they are written centered-worlds-style, centered worlds will come out.

4) In order to get either Folk or Fancy Functionalist accounts to differentiate between the *de se* and non-*de se* cases, we need to guarantee that the appropriate functionalist clauses themselves be read “in the *de se* way.” Then the functionalism presupposes, rather than illuminates, the essential indexicality.

5) We are very much engaged in some rational reconstruction in this chapter. Appeals to functionalism typically happen at a high...
level of abstraction, and it is as a result difficult to see how specific conclusions for the theory of content will result from the functionalism. We’ve done our best here to sketch what we see as the possible ways to go in descending from abstraction to particulars, and argued that none of these ways succeed. But we don’t claim to have a general argument that there’s no better way to develop the functionalist account. If what we say here triggers more detailed and sophisticated accounts from Lewisians or other functionalists, that would helpfully move the debate forward.
Chapter 7

Indexicality and Immunity to Error

Many philosophers have suggested that (some or all) first-person beliefs have an important epistemological status. In The Blue Book, Wittgenstein says:

There are two different cases in the use of the word “I” (or “my”) which I might call “the use as object” and “the use as subject”....One can point to the difference between these two categories by saying: The cases of the first category involve the recognition of a particular person, and there is in these cases the possibility of an error...It is possible that, say in an accident, I should feel a pain in my arm, see a broken arm at my side, and think it is mine, when really it is my neighbor’s....On the other hand, there is no question of recognizing a person when I say I have toothache. To ask “are you sure it’s you who have pains” would be nonsensical. (1958, pp. 66–67)

The same thought is also found in Shoemaker:

The statement “I feel pain” is not subject to error through misidentification: it cannot happen that I am mistaken in saying “I feel pain” because, although I do know of someone that feels pain, I am mistaken in thinking that person to be myself....My use of the word “I” as the subject of my statement is not due to my having identified as myself something of which I know, or believe, or wish to say, that the predicate of my statement applies to it. (1968, p. 557)

While Wittgenstein’s and Shoemaker’s particular examples involve first-person beliefs, the mere existence of some cases of first-person beliefs with this epistemic property is of course insufficient to establish any sort of interesting Essential Indexicality thesis in epistemology. If Wittgenstein’s belief that Russell has a toothache is (in the right cases)
also immune to error through misidentification, then we seem to have an epistemological classification that cuts across the indexical/non-indexical divide, and hence a classification that cannot be used to argue for what we in Chapter 1 called “EI: Epistemology.”

Consideration of cases has led many to think that there is something distinctively indexical involved in immunity to error through misidentification, or “IEM.” It can seem that Wittgenstein can easily form the judgment that Russell has a toothache based on evidence that is in fact about Moore’s toothache (by seeing Moore at a distance wince and hold his jaw, and mistaking Moore for Russell), but cannot form the judgment that he himself has a toothache based on evidence that is in fact about Moore’s toothache. And many have gone on to endorse a theoretical perspective on IEM that makes it either distinctively first-person or more broadly distinctively indexical. Thus Recanati says:

One way of understanding identification-freedom is as follows. Immunity to error through misidentification characterizes thoughts that are “implicitly” de se, as opposed to thoughts that involve an explicit self-identification. Thoughts that are implicitly de se involve no reference to the self at the level of content: what makes them de se is simply the fact that the content of the thought is evaluated with respect to the thinking subject….We end up with a threefold taxonomy. First, there are accidental de se thoughts, namely thoughts that are about an individual x who happens to be oneself. Second, there are genuine de se thoughts—the sort of thought one expresses by using the first person and which, for that reason, I call “first-person thoughts”. In this category we must distinguish between explicit and implicit de se thoughts. Explicit de se thoughts are a variety of de re thoughts and give rise to Frege cases, while implicit de se thoughts are not and do not. (2007a, pp. 4–5)¹

For Recanati, IEM is restricted to (and co-extensive with) the implicitly de se, and is not a feature of just any de re thoughts. Evans ties IEM to demonstrative thought more broadly:

It is a consequence of what I have said about “this”-Ideas and “here”-Ideas that they give rise to the possibility of knowledge which is identification-free

¹ See also his 2007b, esp. pp. 147–148.
in the narrow sense....Clearly, judgments involving Ideas which give rise to the possibility of identification-free knowledge have an epistemological priority over judgments involving all other Ideas of objects: were there no such judgements or Ideas, no singular knowledge would be possible....The contrast between identification-free and identification-dependent knowledge...comes out very clearly in the case of propositions involving a “this”-Idea. If a subject has such an Idea, he will be receiving information from an object in some modality—along some channel. If he knows, or takes himself to know, upon the basis of the operation of the channel, that the property of being red is instantiated, then he may judge “This one is red”, and such a judgment is identification-free....Since [such judgments] do not rest upon an identification, they are immune to error through misidentification. (1982, pp. 181–182)

Evans notes that Shoemaker and Strawson “both...restrict the point—wrongly, as I shall argue—to ‘I’-thoughts (and to a special sub-class of ‘I’-thoughts at that)” (p. 182 fn. 55).

We will argue for three theses regarding this proposed epistemic route to Essential Indexicality. First, there is in fact no philosophically interesting phenomenon of immunity to error through misidentification. Second, to the extent that there is an identifiable phenomenon at all, it is not deeply or constitutively an indexical phenomenon—there are numerous examples of non-indexical beliefs that are immune to error through misidentification. And third, even if our examples of non-indexical immunity are unpersuasive, there is no plausible story about the source of the epistemic status of IEM beliefs that derives the epistemic status from any kind of indexical content, so the distinctiveness would not, in any case, be an argument for EI: Epistemic as we characterized it in Section 1.5—namely, as a thesis about the explanatory power of some level of indexical content.2

2 Furthermore, we should expect that there are, somewhere, non-indexical instances of immunity (even if our cases are not those instances), since systematic indexicality of IEM cases would be a mystery absent a story that derived immunity from indexical-specific sources.
7.1. Thesis 1: There Is No Immunity to Error Through Misidentification

Consider a paradigm instance of supposed immunity to error through misidentification, such as Ludwig’s claim, on the basis of his painful phenomenology:

(*) I am in pain.³

What exactly is it that is being claimed about (*) when it is called “immune to error through misidentification?” Some characterizations are clearly implausible. Wittgenstein in the earlier quotation from The Blue Book says that it would be “nonsensical” to ask whether (in this case) it was really you, rather than someone else, who is in pain. But of course it is not, in the strict sense, nonsensical. The question is clearly sensible; its sensibility is what makes it so clear that the answer is “yes.” (Similarly, many people say that in cases like (*), it “makes no sense” for Ludwig to wonder who is in pain. But again of course it is perfectly sensible to wonder, even if it is epistemically ill-advised or pointless.)

The more interesting claim is that (*) is epistemically privileged, by virtue of a certain kind of error being impossible, in the sense of a priori ruled out. This claim can be left unprecisified (there is some way of going wrong that the believer can be certain not to have occurred), or given a much more precise analysis as in Pryor’s definition of immunity to error through wh-misidentification, according to which:

A singular proposition about a to the effect that it is F is **vulnerable to wh-misidentification** when justified by grounds G just in case:

When that proposition is believed on grounds G, it is possible for those grounds to be defeated by undercutting evidence in such a way that the following two conditions hold:

(i) the combination of G and that undercutting evidence no longer justifies you in believing of a that it is F; but

³ Some proponents of IEM claim that IEM is limited to first-person psychological attributions, while others extend the IEM phenomenon to first-person physical attributions as well. Some proponents of IEM claim that all beliefs with the right sort of content are IEM, while others claim that contents are IEM only relative to an epistemic ground. Our paradigm instance conforms to the more restrictive criteria in both cases.
(ii) the combination of G and that undercutting evidence could, by itself, offer you knowledge that $\exists x \ F(x)$. (1999, p. 284)

Here is our central argument for Thesis 1: no claim, including (*), is epistemically privileged in this way. Suppose Gareth believes that his legs are crossed on the basis of proprioceptive awareness. He could, of course, be wrong about the state of his legs (his proprioception could be malfunctioning), but the received wisdom is that he could not be wrong about whose legs are (purportedly) crossed. But suppose Gareth has, through some science-fiction mechanism, been wired up so that he sometimes receives proprioceptive input from John’s legs. Then, when Gareth has proprioceptive awareness as of his legs being crossed, he should be, if he is aware of his deviant wiring, uncertain whether it is his legs or John’s legs that are being proprioceptively represented as being crossed. So even if Gareth has not been deviantly wired, he is not immune to error, because until the possibility of deviant wiring has been eliminated, the possibility of error due to misidentification cannot be eliminated. There is clearly a general recipe here: for any of Gareth’s beliefs about himself and any source of evidence for that belief, we can imagine him being such that he sometimes receives evidence of that sort from someone else’s state. 4

Cases of this sort have been much discussed in the literature (Higginbotham’s “Davidson’s Eyes” cases, cases of schizophrenic “thought insertion,” Evans’ “wild shot in the dark” (Higginbotham 2010, Campbell 1999, Evans 1982)). But no satisfactory account has been given of a philosophically interesting epistemic phenomenon that survives the problem cases. Sometimes it is suggested that when Gareth is deviantly wired, he thereby comes to own the states and objects he is deviantly aware of, and hence does not fall into an error of misidentification if he identifies John’s legs, or John’s pains, or John’s happiness, as his own. This line is wildly implausible in the case of legs, but even in the case of mental states such as pains or happiness, it

4 For example, maybe what is IEM for Gareth is his belief that his legs seem to be crossed. But now imagine that Gareth is deviantly wired, so that he receives “introspective” awareness from John’s seemings. Then when Gareth finds himself believing that his legs seem to be crossed, he can plausibly wonder whether it’s really his seeming, or John’s. Thanks to Jason Schukraft for encouraging us to address this sort of case.
remains implausible. Gareth can stand in an epistemic relation to John’s pains (such as being aware of those pains, or being reliably disposed to form beliefs about those pains) without having all of the relations to mental states characteristic of ownership.

A more plausible line of response originates in Evans. Evans suggests that in these “Davidson’s Eyes” cases, the subject is not in a position to know anything, including the existential belief that someone has his legs crossed (etc.), and hence that there is no danger posed to IEM—no threat that one forms a belief that is not about the person about whom belief is justified by the evidence, because belief about no one is justified by the evidence. Evans suggests that in “deviant wiring” cases, any belief would be a “shot in the dark”:

We cannot think of the kinaesthetic and proprioceptive system as gaining knowledge of truths about the condition of a body that leaves the question of the identity of the body open. If the subject does not know that he has his legs bent (say) on this basis...then he does not know anything on this basis. (To judge that someone has his legs bent would be a wild shot in the dark.) (1982, p. 221)

But for Evans’s strategy to be successful, we need a story about the justification relation between (for example) proprioceptive states and beliefs about bodily configuration which allows for justification of the de re belief in the good cases, but does not allow for justification even of the existentially quantified general belief in the bad cases. No such story is given by Evans. And there is general reason to think that no such story is possible. If proprioceptive states have content (on the basis of which they justify beliefs), then that content is either general or singular. If it is general, its ability to justify general beliefs should be unimpaired by deviant wiring. If it is singular, then it may fail to justify the existential belief if its singular content is inevitably and invariably first-person and if discovery of “deviant wiring” then uncovers that the causal origin of the perceptual state may not match its content. But that means the guarantee of immunity is only as strong as the guarantee that proprioception (for example) only ever represents the proprioceving agent. Even if it’s true that proprioception functions in humans in this way, it doesn’t look epistemically guaranteed. (The very conceivability of deviant wiring cases shows this.)
For IEM to be a true form of *immunity*, then, we would need to be able to rule out deviant-wiring cases in advance. We cannot. So the most that can be said here is that under normal circumstances, when Ludwig forms the belief (in the normal way) that he has a toothache, he is not wrong about who has a toothache. But that is to say no more than that our beliefs, in normal circumstances, track the world. There remains nothing philosophically distinctive about the specific *immunity* thesis.

### 7.2. Thesis 2: Immunity Does Not Track Indexicality

Suppose Thesis 1 is false, and that there is a philosophically interesting epistemic status that some beliefs have (on some grounds) of being IEM. That status is relevant to the current project only if it is distinctively characteristic of indexical beliefs. But there is no good reason to think that IEM is distinctively characteristic of indexical beliefs, and ample reason to doubt that it is.

Direct arguments that there is a constitutive link between indexicality and IEM are vanishingly rare in the literature. Instead, the link between indexicality and IEM is simply driven by the fact that the typical examples of immunity involve indexical expressions, and the typical examples of *vulnerability* to error through misidentification involve non-indexical expressions (such as proper names). This kind of quasi-inductive reasoning is, of course, extremely vulnerable to the exact distribution of the data. Once we notice that not all first-person beliefs are IEM, but only (for example) first-person beliefs involving certain predicates (“is in pain,” “has crossed legs,” but not “is in Paris”) and justified by certain forms of evidence (introspective, low-level perceptual, but not testimonial), the two categories no longer seem to fit together so neatly. And when we notice that many other indexical

---

5 It is possible to extract an attempt at such an argument from some of the specific theories of indexical content that we consider in our discussion of Thesis 3, but the failure of these theories to provide an adequate account of IEM means that any direct argument from them will fail to secure the desired constitutive link.
expressions seem to have no particular connection with immunity ("tomorrow," "local" (if we think that is indexical)), the inductive evidence is further weakened. Indexicality is thus not sufficient for immunity. Neither is it necessary. Campbell (1999) has observed that descriptive names create immunity. If Holmes believes that Jack the Ripper is the killer, he may be wrong, for there may be no killer. But he cannot be wrong because he is thinking about the wrong person: he is thinking about the killer, if there is one, or about no one, if there is not. Descriptive names are not indexical, so here we have immunity without indexicality.

Even in the standard cases of IEM some further argument is needed to connect IEM with indexical content. Suppose that Ludwig cannot be subject to error through misidentification when he attributes a toothache to himself, and Russell cannot be subject to error through misidentification when he attributes a toothache to himself, and so on. It does not yet follow that indexicality is involved. Perhaps Ludwig has epistemic capacities that are particularly attuned to the state of Ludwig, Russell has epistemic capacities that are particularly attuned to the state of Russell, and so on. Stating this distribution of capacities might most naturally be done in indexical language, but the capacities, and resulting beliefs, need not themselves be indexical-involving. Perhaps what is IEM for Ludwig is his belief that Ludwig has a toothache, when arrived at in the right way.

The same point holds with respect to demonstrative beliefs. Suppose it is true that when an object is in Gareth’s visual field, and he thinks of it that it is $F$, he cannot be in error by virtue of thinking about the wrong object. Perhaps this is because Gareth is believing that that is $F$, and the demonstrative character of the thought is crucial to the IEM status. But perhaps he is believing of $o$ that $o$ is $F$, and the IEM status is secured by the intimacy of the visual justification route in this particular case. No demonstrative concept seems necessary to the story (indeed, the difference between Gareth’s using a demonstrative concept and Gareth’s using a non-indexical ready-created de re concept is hard to pin down in the first place).

The appearance of indexical involvement in immunity cases is heightened by a bias in the way the cases are described. Suppose Gareth knows the position of his legs via proprioception. One way of
characterizing the situation is by saying that Gareth knows “My legs are crossed,” and cannot be wrong about whose legs he believes to be crossed. But the situation need not be characterized first-personally. Call Gareth’s two legs “LL” and “RL.” Another way of characterizing the situation is by saying that Gareth knows that LL is on top of RL, and cannot be wrong about which two legs he is thinking about. On this way of thinking about things, proprioception is not focused single-mindedly on the self, but is focused on a number of objects—the component bodily parts that belong to the self. There is no obvious need for a concept of the self, or for any indexicality, to be involved. It is the theoretical choice to describe proprioception as delivering self beliefs, rather than LL, RL, etc. beliefs, that makes IEM look indexical-involving.

7.3. Thesis 3: Indexicality Does Not Explain Immunity

Suppose our first two theses are incorrect: immunity to error through misidentification is a real phenomenon, and we have not convincingly argued for non-indexical cases of immunity. Essential Indexicality in the sense that we are concerned with does not yet follow. Our concern is with the claim that indexicality of (some level of) content plays special philosophical roles. Even without Theses 1 or 2, all we have is that (a) certain beliefs, justified in certain ways, have a certain special epistemic status, and (b) the IEM beliefs we have considered systematically have indexical contents (bracketing for the moment how one should best characterize that content). It does not follow that the indexical contents play an explanatory role in the epistemic status—both could, for example, be consequences of a common cause (such as the causal proximity of the object of the belief).

What is needed, then, is an account that shows how IEM flows from indexical contents. But our third thesis is that no such story succeeds. We consider three accounts: the role of identification-free justification (along the lines that Evans suggests), the use of relativistic/centered-worlds contents that only implicitly represent the self (following Recanati’s proposal), and Higginbotham’s appeal to
token-reflexive contents. These stories all either fail to explain immunity, or explain it in a way that has no connection to indexicality.

**Story 1: Identification-Free Justification**

IEM is a consequence of beliefs whose justification does not include any identity claims. Contrast two cases:

**(No Immunity)** Gareth believes that *that is* F, and that *that* is identical to John, and thereby comes to believe that John is F. But Gareth’s belief is subject to error through misidentification: he could be right that someone is F, but be wrong that it is John that is F. This will happen if the identity claim in his justification route is incorrect.

**(Immunity)** Gareth believes, on the basis of proprioception, that he himself is F. The belief is not justified by way of an identity between a first and second presentation of Gareth, and hence is immune to error through misidentification.

The identification-free justification strategy fails in two ways. First, identification-free justification is not in fact sufficient to achieve immunity to error through justification. It is true that if there is a deductive justification route for the belief that does not appeal to any identity premises, then no false identity beliefs on the part of the believer can endanger the justification, simply because of the monotonicity of deductive consequence. But if the justification route is inductive, then false identity beliefs not occurring in the justification can undermine the justificatory status, potentially removing the object-dependent justification while preserving justification for the existential belief. Since very many beliefs, including purportedly IEM ones, are justified inductively rather than deductively, the identification-free status does not deliver the desired account of immunity.

Second, even if the identification-free strategy had succeeded, it would not have shown that the epistemic phenomenon of immunity was rooted in indexicality. Identification-free justification routes are not peculiarly indexical beasts. Gareth begins with a perceptual experience

---

6 At least, no one has even claimed, let alone established, that all IEM beliefs are purely deductively justified.
with the content that John is $F$, and that experience then justifies Gareth’s belief that John is $F$. No identity claim plays a role in the justification—Gareth simply moves directly from the content of the perceptual experience to the corresponding belief. Thus immunity is guaranteed, and immunity is not especially associated with indexicality.7 Or, if one thinks that perceptual experiences cannot contain contents like “John,” Gareth’s perceptual experience can have the content $o$ is $F$, and that content can stand in a primitive justificatory relation (perhaps grounded in reliable perception-to-belief transition mechanisms) to Gareth’s belief that John is $F$. Again, we have a case of identification-free belief justification without indexicality.

**Story 2: Relativistic/Centered-Worlds Contents That Only Implicitly Represent the Self**

Recanati argues that some first-person mental states—in particular, those yielded by low-level epistemic faculties like proprioception and introspection—involve what he calls the “implicit de se.” In such states, there is no semantic representation of the self—the semantic content is thus something along the lines of properties or sets of centered worlds. The states receive their truth conditions in terms of the self, because it is a feature of the mental-representation system as a whole that a representational state whose content is a property is veridical if and only if the representing agent possesses that property. Because there is no representation of the subject, there is no danger of misrepresentation of the subject.

In light of our earlier remarks on the role of self-ascription in Lewis’s account of the de se in Chapter 5, our objection to Recanati’s strategy should be predictable. If Recanati’s appeal to implicit representation

---

7 We are granting at this point the falsity of Theses 1 and 2, to show that even without them, Thesis 3 suffices to defeat EI: Epistemology. In granting the falsity of Thesis 2, we are granting that we have not exhibited non-indexical instances of immunity to error through misidentification. Granting that, however, is consistent with the possibility of there being such non-indexical cases—we grant only that we haven’t yet given positive reason to endorse the claim that there are. But from the observation that identification-free justification routes are not specifically indexical, together with the proposed claim that identification-free justification routes ground IEM, it looks like it will now follow that there are non-indexical cases of immunity.
is successful, it is the implicitness that does all of the work. What matters for securing immunity is that the subject of the belief not be explicitly represented in the belief state. But the use of property-like representations is compatible with any number of background stories about the metasemantics of those representations. Recanati assumes that it is always the function of the representing system to deploy property-like representations to represent the (merely implicit) self (to “self-ascribe” those properties). But one could just as well have a representing system that deployed property-like representations to represent (merely implicit) Gareth. Again immunity would result (if Recanati’s strategy works), but without indexicality. The indexicality of Recanati’s story is located entirely in the way in which the global system deploys the implicit representations, but that part of the story is irrelevant to achieving immunity.

*Story 3: Appeal to Token-Reflexive Contents*

Higginbotham (2003, 2010) has argued that immunity to error through misidentification is a consequence of special token-reflexive contents. When Sydney remembers answering the question, for example, he undergoes some token mental state $e$. The content of state $e$ then contains a reference to $e$ itself, which is (roughly): the experiencer of $e$ answered the question. Sydney is thus in no danger of misidentifying the subject of the remembering, because the subject is given to Sydney as the experiencer of the very mental event that Sydney is at that moment undergoing, and Sydney cannot be uncertain who that is.

But all Higginbotham’s token-reflexive contents do is to place a particular object in the content of the mental state, as a *de re* constituent of it. If Higginbotham’s strategy is to succeed, there must be no danger of Sydney’s misidentifying not *himself*, but the mental event $e$. But surely there is such a danger—Sydney might be exposed to $e$ not “from the inside” (whatever that means), but by (for example) seeing $e$ on a brain scan. If there is no danger of misidentification of $e$, this is

---

8 We are not yet convinced that it is, since the distinction between representing by way of a local feature of the representational state and representing by way of a global feature of the representation-supporting system has yet to be developed in a convincing way.
not a result of a (further) token-reflexive content, but an unexplained further immunity fact. So, at bottom, immunity for Higginbotham is not explained by indexicality, but is effectively posited as a basic epistemic relation to objects.

7.4. Brief Summary of This Chapter

This chapter has three rather simple take-home messages:

1) There is no such thing as immunity to error through misidentification. As a consequence, the thought that indexicality is philosophically important because it gives rise to IEM has to be wrong.

2) If you think we are wrong about (1), then we can exhibit cases of immunity to error through misidentification that do not involve indexicality, and cases of vulnerability to error through misidentification that do involve indexicality. So even if there is such a phenomenon, it isn’t attached specifically to indexical states.

3) If you think we are wrong about (2), we note that in none of the theories proposed for how to secure IEM does indexical content play an important explanatory role. You can think of that point in two ways: (i) as a matter of fact, no one has proposed a positive theory of IEM that denies (2), (ii) if we construe extant theories as attempts to deny (2), they are internally inconsistent. In either case, we are left without an account of IEM that underwrites EI: Epistemology.
Chapter 8

A Brief Note on Perceptual Content and the De Se

Pre-theoretically it is tempting to treat the input and output sides of contents analogously: if there are irreducible de se beliefs, expectations, assertions, etc., then we should expect the input that gives rise to these outputs, i.e. perception, to also contain (or at least be characterized by) such contents. In short, you might be tempted by this thought:

(T) if there’s irreducible de se (or other indexical) content anywhere, it is in perception.

Suppose all of your perceptual contents were objective and non-perspectival, but that the beliefs you formed on the basis of those perceptions were frequently perspectival and de se. Prima facie, this plays havoc with the epistemology: how are your perceptual states going to justify beliefs which systematically turn out to be about something quite different from what is perceptually presented? That pre-theoretic thought triggers three questions:

1) Are there arguments in the recent literature on perceptual content that aim to establish some version of the view that such contents are essentially indexical?
2) If the answer to (1) is “yes,” do these arguments add anything to what we have considered in previous chapters? i.e. suppose you agree with all of the anti-EI arguments presented in previous chapters, is there anything in this literature that can sway you in favor of EI?
3) If the answer to (1) or (2) is “no,” what general lessons about EI can we draw from that?
In the remainder of this brief chapter, we show that the answer to (1) is a qualified “no”: it is widely assumed that there are such arguments, but they are hard to find in published work. We then show that the arguments that are standardly gestured at are the same as those we have already discussed—they don’t add anything to what we have already considered. In other words, if you have been swayed by the arguments in previous chapters, then there’s simply nothing new in this literature. At the end of this chapter, we draw some more general lessons.

8.1. No New Arguments for De Se

Contents

Regarding question (1): there are surprisingly few explicit endorsements of essential indexicality (or irreducible de se contents) in the recent literature on the contents of perception. That should worry proponents of irreducible de se contents. (T) makes available a modus tollens argument. Given that if there’s irreducible de se content anywhere, it is in perception, if we also have there’s no irreducible de se content in perception, it will follow that there’s no irreducible de se content anywhere. So if there’s no argument for essential indexicality in perception, there’s a threat to the entire essential indexicality project. Some prominent writers, e.g. David Lewis, have noticed this. The assumption that perception is irreducibly de se is at the center of one of Lewis’s most important arguments in his 1979 paper. Discussing the case of Lingens the amnesiac—one of the central motivating cases in the paper—Lewis says:

Book learning will help, no doubt, but only because Lingens has more than book learning. He is in a position to self-ascribe the property of being in a certain perceptual situation. This is a property that does not correspond to any proposition, since there are worlds where some have it and others do not. Book learning may eventually convince Lingens that he inhabits a world where exactly one person is in that perceptual situation, and where that one is Rudolf Lingens, who is in aisle five, floor six, of Main Library, Stanford. Then his problem is solved. But not because he has managed to conjure nonpropositional belief out of propositional belief. He relied on his perceptual belief, and that was already nonpropositional. Nonpropositional plus propositional
belief can give more nonpropositional belief. That is how Lingens can find out who and where in the world he is. (p. 520)

So as Lingens, the amnesiac, reads more and more, he learns that there is this man, Lingens, who has such-and-such perceptual inputs, e.g. as of reading books about a man in the Stanford Library. Putting all that together with his perceptual input, he concludes that he is Lingens (basically he thinks: this Lingens guy has all the same perceptual inputs as me, so I must be Lingens). Didn’t he solve his problem by reading books? No, says, Lewis, “He relied on his perceptual belief, and that was already nonpropositional.” Now, Lewis never tells us why perceptual beliefs are non-propositional in his sense. In the remainder of this chapter we argue that Lewis was mistaken.¹

Susanna Siegel usefully summarizes some of the considerations that speak in favor of that view. She says:²

Many contentful experiences seem to have contents that must be specified by the use of *indexical* expressions, such as “over there”, “to the left”, “here”, “in front of/behind me”, “just a second ago”, “since a few seconds ago”, and so on. For instance, some auditory experiences seem to present sounds as coming from a direction relative to the ears; some proprioceptive experiences seem to present pressure sensations as located in specific parts of the body; visual experiences seem to present things in locations relative to the eyes and the rest of the body. More generally, contentful experiences present the world from the subject’s perspective both in space and in time. If one tried to characterize the perspective aspect of contents in language, it would be natural to use indexical expressions for spatial properties, for the subject herself, and for times. (2011)

Two points are important for our purposes. First, Siegel points out that we naturally describe perceptual contents using indexicals: *I see the tree over there / in front of me.* This much is indisputable. Siegel is surely right also to point out that when one tries to characterize the

¹ Maybe we are reading too much into the passage from Lewis; after all, Lewis only says that Lingens relies on his non-propositional perceptual beliefs, and doesn’t explicitly say that the perceptual content is non-propositional. However, Lewis (1980) asserts that “the content of the [perceptual] experience is, roughly, the content of the belief it tends to produce” (p. 239), so we think it is reasonable to extend Lewis’s talk here from the non-propositional perceptual beliefs to non-propositional perceptual contents.

² This is from her *Stanford Encyclopedia* entry on perceptual contents—and so shouldn’t be taken to represent Siegel’s view, but rather her view of the current consensus in the literature.
perspectival aspect of perceptual contents in language, it is natural to do so using indexicals. We emphasize “natural” in the previous sentence to emphasize the contrast with what Siegel says in the first sentence of the quoted passage: “Many contentful experiences...must be specified by the use of indexical expressions” (bolded emphasis added). It is this latter claim we want to object to. It is beyond dispute that using indexicals is natural, but, if our view is right, it is never necessary. Whatever can be expressed indexically could be expressed by non-indexical means. So what we are looking for (and we expected to find when studying this literature) are arguments for the stronger claim—the claim at the start of the passage from Siegel. It is useful to look at the arguments sketched by Siegel right after the passage quoted:

To begin with spatial content, suppose you hear a sound as coming from your left. Someone else facing you, due to different orientation, hears the same sound as coming from the right. Moreover, a third perceiver situated halfway around the world could hear a qualitatively identical sound as coming from her left. The area to the first perceiver’s left is the very same spot as the area to the second perceiver’s right, whereas the area to the left of the third perceiver, who is very far away from the other two, is a completely different part of space. Now, if contents are to reflect which side of the perceiver the sound seems to come from, then it is not enough simply to include the very area from which sounds are coming in the content. That would not reflect what’s in common between you and the third perceiver, who is halfway around the world. It would also fail to reflect what’s different between you and the second perceiver, who is differently oriented than you. So it seems that some more specific representation of location and direction is needed.

To make the case a bit more vivid, let the three agents be Ida, Sophia, and Nora. Let’s assume that “x and y hear S” means that x and y hear a qualitatively identical sound, S. These are the stipulated facts:

• Nora hears sound S as coming from her left.
• Sophia, who is facing Nora, hears sound S as coming from her right.
• Ida, who is in a different part of space from Nora and Sophia, hears sound S as coming from her right.
About this scenario it is plausible to say the following:

**Sameness.** There’s something important in common between Ida and Nora (pre-theoretically, that they hear \( S \) as coming from the *same* direction.)

**Difference.** There’s some important difference between Nora and Sophia (pre-theoretically, that they hear \( S \) as coming from *different* directions).

If the three of them try to express or describe the contents of their experiences in language, the use of indexicals would be *natural*, but surely not *necessary*. Nora could say, “Nora hears sound \( S \) as coming from Nora’s left.” Sophia could say, “Sophia hears sound \( S \) as coming from Sophia’s right.” This satisfies Difference since only one of them (Nora) is an \( x \) such that \( x \) hears \( S \) as coming from \( x \)’s left. The other one (Sophia) is an \( x \) such that \( x \) hears \( S \) as coming from \( x \)’s right. In other words, accounting for Difference is easy and there’s no need to use any kind of indexical expression.

The same is true of Sameness. Ida can describe the content of her experience by saying “Ida hears \( S \) as coming from Ida’s right” and Sophia can say, “Sophia hears \( S \) as coming from Sophia’s right.” As a result, there’s something Sophia and Ida have in common: each is an \( x \) such that \( x \) hears \( S \) as coming from \( x \)’s right. Again, there is no need to appeal to any kind of indexicality.\(^3\)

As we have set things up, Sameness and Difference are captured in terms of quantified principles governing relations between perceivers and contents. Ida and Sophia do not have the *very same* perceptual content, but they stand in the same sort of relation to the contents they have. But even if someone insists that Sameness must be captured via exact identity in content (and that the abstractions over contents via

\(^{3}\) Those familiar with the recent debate between contextualists and relativists will notice that the moves we have just made are analogous to moves made by contextualists in response to relativist accusations that they fail to account for sameness and difference in content of utterance of, e.g. “\( X \) is delicious” (see, e.g. the discussion in chs. 2 and 4 of Cappelen and Hawthorne 2009). One point of disanalogy is worth noting: the appeal to so-called eavesdroppers is one of the powerful considerations in favor of relativism. Note that no analogous move is available in this debate: there is only one perceiver and we are focusing on the content of that perceiver’s input.
which we capture Sameness and Difference are not deserving of the name “content”), there is still no argument for Essential Indexicality here. A content that is shared by both Sophia and Ida need not be a deeply de se content. There is no need here for a first-person concept, or for a (Lewisian) picture of self-ascription of the perceptual content.

To see this, suppose Maeve has a non-standard perceptual system, which represents how things are around Brigid, instead of around her (perhaps via a camera attached to Brigid feeding a signal to Maeve’s brain). Maeve has an auditory experience as a consequence of a sound S to the right of Brigid. Maeve’s auditory experience is then phenomenologically indistinguishable from those of Sophia and Ida, and hence (according to the proponent of the stringent Sameness requirement) has the same content. But that content is not self-ascribed. Its truth conditions are not in terms of the perceiver (Maeve), but in terms of Brigid. Once the point is seen, we can collapse Maeve and Brigid into a single perceiver, and see that there is no need for a distinctively first-person feature of the perceptual content.

We take it that we are in no way disagreeing with Siegel. She summarizes these points as follows:

If locations and directions are represented in experience, then it seems they are represented relative to the subject, in which case there has to be some way of representing the subject herself (or itself)—or at least the perceptual organs or body parts that are on the subject-side of the spatial relations.

What we want to draw attention to here is simply that no argument has been given that representing a subject or a body part or a perceptual organ must be done indexically or in a distinctly de se way. While it might in some cases be natural to do so, none of these arguments (or the data appealed to) even purport to show that it is required.

At the beginning of this chapter we asked two questions:

1) Are there arguments in the recent literature on perceptual content that aim to establish some version of the view that such contents are essentially indexical?
2) If the answer to (1) is “yes,” do these arguments add anything to what we have considered in previous chapters? i.e. suppose you agree with all the anti-EI arguments presented in previous
chapters, is there anything in this literature that can sway you in favor of EI?

The answer to (1) is: surprisingly few—there are some statements of that view in, e.g. Lewis and Siegel, but very few attempts to explicitly argue for it. In general, over the last thirty–forty years there is a surprising lack of connection between the literature on the contents of perception and the literature on essential indexicality. In reply to (2), the answer is also “no.” The arguments we do find, in so far as we find any, are really just versions of arguments we have seen earlier, in connection with Lewis on information gathering, Perry on Lingens, Perry on agency, etc. So our overall conclusion is that if you are sympathetic to our anti–EI line in previous chapters, you will find nothing in the perception literature to change your mind. That brings us to our third question: What general lessons about EI can we draw from this? We’re somewhat moved by the *modus tollens* argument mentioned earlier: *If there’s irreducible de se (or other indexical) content anywhere, it is in perception. There isn’t any in perception (no one has even tried to show that there is). So, it’s nowhere.*

8.2. Addendum: Brogaard and Chalmers on Perceptual Content

We should note that Chalmers (2006) and Brogaard (2009) have argued that these perceptual properties can’t be relational (in the way that we have proposed here), because their phenomenology reveals them as monadic. Brogaard, for example, endorses a view she calls “Weak Representationalism”:

...the view that the phenomenology of perceptual experience determines the content of perceptual experience. On this view, a phenomenally red experience has a content that contains a red color constituent. Given its weakest formulation, representationalism is exceedingly plausible (Tye 2000, Chalmers 2004, Siegel forthcoming). (2009, pp. 221–222)4

4 It is worth noting that in personal correspondence, both Tye and Siegel deny ever having endorsed what Brogaard calls “Weak Representationalism.”
She says that Weak Representationalism implies that we shouldn’t endorse the account of Sameness and Difference that we have given since the content of the relevant experiences shouldn’t “contain the subject.” That, it is claimed, is phenomenologically implausible. Here is Brogaard’s illustration of the point: 5

When I look at two trees at different distances from me, I can see that one tree is further away from me than the other. Moreover, it is plausible that you and I can have perceptual experiences with the same phenomenology of the two trees (perhaps at different times). By weak representationalism, our experiences have the same content. So, our perceptual experiences cannot contain you or me in the content of perception. (2009, p. 223)

Similarly, Chalmers endorses a principle of Phenomenal Revelation—that the phenomenology of perceptual experience reveals something about the logical form of the represented features. Chalmers (2006) asserts that “the properties attributed by color experiences are non-relational properties” and justifies this assertion simply by claiming that “[it] is grounded in the phenomenology of color experience” (p. 56).

There are two related arguments in this vicinity, both of which are irrelevant to issues having to do with EI.

**Argument from Weak Revelation:** When Sophia and Ida both hear sound $S$ to the right, they both have auditory experiences with the same phenomenology. The non-indexical explanation of the perceptual sameness between Sophia and Ida was in terms of a quantified principle—while Sophia had a perceptual state with the content *there is a sound to Sophia’s right* and Ida had a perceptual state with the content *there is a sound to Ida’s right*, there was a sameness between them by virtue of their both *being an $x$ such that $x$ has a perceptual state with the content that there is a sound to $x$’s right*. This explanation, however, is incompatible with Weak Representationalism. Given Weak Representationalism plus the phenomenological sameness of Sophia and Ida, it then follows that their auditory experiences have the same content, contrary to what the non-indexical explanation requires.

---

5 What is in effect the same argument can be found in Chalmers 2006.
Argument from Phenomenal Revelation: The non-indexical explanation of the perceptual sameness of Sophia and Ida requires making the content of their perceptual states relational, rather than monadic. Sophia represents S as being to the right of Sophia, rather than just to the right. But by Phenomenal Revelation, the phenomenology of the perceptual experience simply reveals the represented property as monadic, rather than relational, ruling out the non-indexical explanation.

Suppose you find these arguments convincing. What follows? It should be obvious that nothing follows about indexicality (and that is why we have put this discussion in an addendum). As we have emphasized, what is important here is not whether we place the perceiving subject “in the content” or not, but whether that perceiving subject has to be represented in a distinctly de se or indexical way. Nothing in the arguments here even starts to motivate the latter idea. The arguments, if they succeed, at best tell us something about the location or role of the perceiving agent in the representation of the content of perception. The arguments don’t show, and don’t purport to show, that the agent has to be represented indexically (or in a distinctly de se way).

A possible source of confusion here is that both Brogaard’s and Chalmers’s alternatives to our account of Sameness and Difference appeal to centered-worlds contents; that is, they draw on Lewis’s (1979) semantic framework. Here is Siegel’s succinct summary of this line of thought (this is again from the 2011 Stanford Encyclopedia entry under the heading “Indexical Content”):

An influential approach that has been developed for both beliefs and experiences appeals to a special kind of possible world: centered worlds, where these are possible worlds marked with a designated subject and perhaps a designated time or other parameters. The centered-world framework can provide a level of content that can be shared by two perceivers who perceive qualitatively identical but numerically distinct cubes in different locations as being nearby: each experience is true in worlds where there is a cube nearby the designated center of that world. Features such as being nearby the center are called “centering features” (see Egan 2006a, 2006b, 2010), and can be used to describe both the phenomena that indexicals are apt for describing, as well as perceptual constancies and secondary qualities.
Perceptual Content and the de se

Does the use of centered worlds in characterizing perceptual contents, in itself, provide support for some version of Essential Indexicality? Only insofar as Lewis’s framework in general supports such a thesis, and as we argued in Chapters 5 and 6, it doesn’t. We won’t rehearse those arguments here.\(^6\)

8.3. Brief Summary of This Chapter

Here are three take-away messages from this chapter:

1) Proponents of de se/perspectival content should be more concerned with arguing for the presence of that content in perception than they have been. Given the status of perception as the

\(^6\) Another argument along these same lines can be found in Schroeder and Caplan (2007), who argue that because subjects can be placed in perceptual situations that are phenomenologically indistinguishable but which call for different veridicality conditions, there is a difficulty for accounts that explain sameness of phenomenology via sameness of some level of perceptual content. Again, we note that even if this explanatory aspiration is granted (that is, if Brogaard’s Weak Representationalism is endorsed for some level of content), no interesting form of essential indexicality follows.

We have suggested here that the phenomenal sameness of Ida and Sophia is to be accounted for via a quantified principle relating the bearer of a perceptual content to that content — both Ida and Sophia are an \(x\) such that \(x\) hears sound \(S\) to \(x\)’s right. This quantified principle can be treated as introducing a kind of content—a lambda-extracted content \(\lambda x (\text{there is a sound } S \text{ to } x\text{'s right})\). If someone wants to call this content a “perceptual content”, we see no reason to object. But there is no reason to think that this content is the carrier of the veridicality conditions for the perceptual experiences of Ida or Sophia (the veridicality conditions can be carried by the fully propositional contents \(\text{there is a sound } S \text{ to Ida’s right}\) and \(\text{there is a sound } S \text{ to Sophia’s right}\)). Hence, there is no reason to introduce indexicalized truth-relative contents into the perceptual story.

But even if indexicalized contents were introduced, they would not amount to the introduction of essential indexicality. Again, mere indexicalized truth is not enough. As the example of Maeve and Brigid shows, indexed truth can be interpreted non-indexically — Maeve can have perceptual contents that have no absolute truth conditions, but whose role in Maeve’s perceptual system is to be true relative to Brigid. And one of the lessons of Chapter 5 was that even when the truth-relative contents are to be interpreted relative to the bearer of the content, essential indexicality is still not created. True essential indexicality would, for example, distinguish between the claim that Maeve had indexed-truth perceptual states whose veridicality conditions were true relative to Maeve and the claim that Maeve had indexed-truth perceptual states whose veridicality conditions were truth relative to herself. Only the latter would have the de se element characteristic of the (putatively) essentially indexical, and nothing in the various arguments from perception imposes a requirement specifically for the latter.
primary “point of entry” of contents into our mental lives, it would be very surprising if mental contents were de se after this point, but not at it. If there is no de se content in perception, it seems likely that there is no de se content anywhere.

2) The primary argument that has been given for perspectival perceptual content has been in order to capture patterns of sameness and difference in perception. But in ways familiar from the contextualism/relativism literature, these patterns can be captured using higher-level abstractions on content. Perspectival content is the representation of relations between the perceiver and the environment, and sameness of perspectival content between $A$ and $B$ amounts to $A$ and $B$ both being an $x$ such that $x$ is so-and-so related to the environment.

3) One can try to block the contextualist response to sameness and difference arguments by appealing to principles strengthening the connection between phenomenology and content such as Weak Representationalism and Phenomenal Revelation. But the attempts to block the move, e.g. from Brogaard and Chalmers, are not arguments for distinctly indexical or de se representations of the perceiving subjects. The alternatives they propose appeal to centered-worlds contents and for reasons given in Chapter 5, such appeals provide no support for any interesting version of Essential Indexicality.
Chapter 9

The De Se and the Semantics of PRO Constructions

This chapter is about PRO constructions. Many philosophers and linguists have claimed that these express a genuine content-level perspectivality. The question we are concerned with in this chapter is whether the semantics of PRO constructions can provide independent evidence (i.e. evidence separable from what we have looked at in previous chapters) of that claim.

The structure of this chapter is a bit complicated. This reflects what we take to be the complicated (and confusing) connection between the linguistics literature on PRO (and related) constructions and the traditional philosophical topics that literature might be relevant to. What is confusing is just how the linguistic data and philosophy connect, if at all. This problem is exasperated by the promiscuous use of theoretical terms such as “de se.” The discussions in philosophy and linguistics have the appearance of continuity since both claim to be about what’s labeled “the de se,” but, as we saw in Chapter 5, this Lewisian term wasn’t well defined by Lewis. Its use in the linguistics literature is parasitic on Lewis’s and so inherits and exasperates the difficulties we found in understanding Lewis’s use. When the term is moved across disciplinary boundaries, the dangers of confusion and obfuscation multiply.

This chapter has two main theses:

1) Even if the data about PRO constructions were all taken at face value, it provides no evidence for anything like an Essential Indexicality thesis and provides no illumination of the philosophical areas where the de se is alleged to be important. This is so for two reasons:
a) It tells us nothing more than what we already knew from reflections about indexicals (and “I” in particular) and the opacity of attitude reports.

b) The mental states described using PRO constructions don’t connect to any of the alleged important philosophical phenomena associated with the *de se*.

2) The PRO data theorists standardly focus on is biased. We show that these constructions can be used in cases where there’s not even a prima facie case to be made for distinctly *de se* (or other perspectival) content. So there is a considerably less close connection between first-person thought and PRO constructions than is typically assumed.¹

9.1. Some Background: With Chierchia the *De Se* Moves From Philosophy to Linguistics

A version of the basic idea we want to challenge in this chapter is first expressed in an early paper by Chierchia. Chierchia (1989) connects the philosophical literature from Castañeda, Perry, and Lewis on essential indexicality with the semantics of certain kinds of attitude reports. His proposal is that the former project will be illuminated by studying the latter. Chierchia (1989) says:

The problem of the so-called “essential indexical” (or in Lewis’s words, of attitudes ‘*de se*’) has been widely discussed in the philosophical literature since it was first pointed out in Castañeda’s work. Various evidence seems to suggest that the perspective one has on oneself, and consequently the attitudes

¹ In connection with (2), some complexities and limitations are worth stressing right away. First, there’s a great deal of non-English data discussed in the linguistics literature (see, e.g. Anand 2006). So one reaction we sometimes get when presenting point (1) is: *Well, yes, the obligatoriness of de se constructions is low in English, but other languages are much stricter about it. You need to look at Amharic (or Ewe, Mandarin, Yoruba, etc.).* We are not going to say much about that additional data, since our main point isn’t affected by it. Our point is simply this: In a lot of the standard cases, the data is much shakier than people have let on.
towards oneself are, in some way, special. The question is to spell out exactly how special. Many proposals have been made in that connection. (p. 3)

Here is how Chierchia describes the initial connection to the semantics of attitude reports:

The question of the nature of attitudes towards oneself, besides being central for cognition in general, has an obvious specifically semantic relevance. From a semantic point of view, the issue is the structure of the logical form (or truth-conditional import) of reports of attitudes towards oneself. For example, one might ask the following question: Can the semantics of such reports be reduced to the semantics of ordinary (i.e. non de se) attitude reports? Or are there special ways in which relations towards self are grammaticized? (p. 3)

There are many constructions now discussed as relevant in this field, but as an introduction to the topic, it will be useful to dwell first on “expects to” (see Morgan 1970, Chierchia 1989). Consider the following scenario:

John is intoxicated and watching TV. He is watching the speeches of various candidates in the upcoming election. He watches one particularly engaging candidate, and comes to think that this candidate will win. The candidate is none other than John himself; but because he is so intoxicated he doesn’t realize that he is the candidate in question. In fact, he is rather pessimistic about his own prospects and thinks to himself, I’m not going to win the election. (Ninan 2010, p. 553)

The judgment at the core of this literature is that the report in (a) is true (or at least has a true reading) and the report in (b) is unambiguously false:

a) John expects that he will win the election.

b) #John expects to win the election.

Why is (b) false even though (a) is true? This is where linguists (and many philosophers) have assumed a connection to the Perry-Lewis literature. Ninan helpfully summarizes the received view:

In order for [(b)] to be true, John would need to think to himself, I am going to win the election. In other words, [(b)] is true iff John has a de se expectation to the effect that he will win. If that’s right, then sentences like [(b)] do not
have coarse truth conditions: they are true only if the subject has the appropriate \textit{de se} attitude. (p. 554)

This is more or less the way, e.g. Anand describes the orthodoxy:

…the subject control form in [(b)] is only compatible with the scenario where John is attributed the \textit{de se} belief expressed with the indexical pronoun \textit{I}, while the ECM control condition [in effect, our (a)] is also compatible with the scenario where John does not have a \textit{de se} belief (2006, p. 15)

An additional assumption in much of this literature, which we will not question here, is that (b) contains a covert unpronounced subject, labeled \textit{PRO}. So the logical form of (b) is something along the lines of (b∗):

\begin{equation}
\text{b∗) John expects PRO to win the election.}
\end{equation}

One standard assumption is that the presence of \textit{PRO} in attitude contexts automatically triggers a “\textit{de se}” requirement.

### 9.2. Some Linguistic Phenomena that Allegedly Require Distinctly \textit{De Se} Interpretations

We start with a brief overview of the kinds of constructions that have been taken to require a distinctly \textit{de se} reading. We already mentioned “expects to” and in this section we outline analogous claims made about “admits to,” “remembers” + gerund, dream reports, and only-binding.

\textit{“Admits to”}

Consider a scenario similar to the one described by Ninan:

John is intoxicated and watching TV. He is watching the speeches of various candidates in the upcoming election. He watches one particularly suspicious candidate, and comes to think that this candidate is a corrupt crook. The candidate is none other than John himself; but because he is so intoxicated he doesn’t realize that he is the candidate in question, he says, “That guy is a corrupt crook”.
As a report of this (c) has a true reading, but (d) does not (we are now assuming that John as a matter of fact is a corrupt crook):

\[
c) \text{ John admitted that John is a corrupt crook.} \\
\#d) \text{ John admitted to being a corrupt crook.}
\]

The alleged reason for the discrepancy is that John didn’t admit of himself \textit{de se} that he was a corrupt crook. The truth of (c) is compatible with John never admitting that \textit{he} is a corrupt crook. Not so for (d). Again, the standard assumption is that the logical form of (d) includes the unpronounced PRO as follows:

\[
\#d^{(*)} \text{ John admitted [PRO to being a corrupt crook]}
\]

\textit{“Remembers” + gerund}

Consider this scenario:

John and Susan visit their struggling graduate student Frank to give him a “shape up” talk. Afterward, John says to Susan, “Did we actually tell him he needs to defend by May?” Susan says, “I remember someone saying it, but I’m not sure which of us.” In fact, Susan said it. (adapted from Higginbotham 2003)

Note that (e) and (f) are acceptable, while (g) is not:

\[
e) \text{ Susan remembers that Susan told Frank to defend by May.} \\
f) \text{ Susan remembers Susan telling Frank to defend by May.} \\
\#g) \text{ Susan remembers telling Frank to defend by May.}
\]

Again, that (g) is unacceptable is taken to be evidence that “remembers” + gerund requires that the agent thinks of herself in a distinctly first-person way, and it patterns with the explanation given of the analogous data for “expects to” and “admits to” since (g) is assumed to have an unpronounced PRO:

\[
\#g^{(*)} \text{ Susan remembers [PRO telling Frank to defend by May].}
\]

\textit{“Only”-binding}

John and Frank are both running for mayor of Springfield, and are both in the local pub in the evening. The evening news shows clips of speeches by both John and Frank. John, rather drunk, fails to recognize
himself in the clip, but is impressed by his speech and concludes that the speaker is likely to win. Frank, sober, does recognize himself in the clip, and is also impressed by his own speech and concludes that he is likely to win. Now consider:

h) Only Frank expects himself to win.
   i) #Only John expects himself to win.

Again, the unacceptability of (i) is taken as to show that distinctly de se contents have to be introduced.

Dream Reports

It has also been claimed that dream-report contexts display a distinctive linguistic pattern which is sensitive to the distinction between de se and non-de se mental states (Percus and Sauerland 2003a, Anand 2006). The putative phenomenon can be illustrated with Lakoff’s (1972) example in which JD utters:

(D) I dreamed I was Brigitte Bardot and I kissed me.

This is an instance of a “counter-identity dream,” or a dream in which I take on an identity other than my own. It’s characteristic of reports of counter-identity dreams that first-person pronouns become ambiguous. They can refer either to my in-dream identity, or to my real-world identity. But the claim is that the dream report (D) is not ambiguous. One might think there were two possible readings:

(R1) In the dream, the BB body kisses the JD body.
(R2) In the dream, the JD body kisses the BB body.

But it is claimed that only the first of these is available. Anand takes this particular case to be an instance of the following generalization:

(BLOCK) No obligatory de se anaphor can be c-commanded by a de re counterpart.

Note that BLOCK, just like the PRO-control explanation context, is a putative instance of a linguistic phenomenon that’s linked to the notion of the de se.
9.3. Our Central Concern: What’s the Philosophical Significance?

Our impression is that it is sometimes assumed that the kind of data just outlined will help us understand distinctively first-person thought better—that an understanding of PRO constructions is important (or at least helpful) for an understanding of distinctly \textit{de se} attitudes.\footnote{If you are someone who doesn’t think that, then this chapter won’t be of much interest and it would probably be a good idea to move on to Ch. 10.} We will be questioning the data itself, but before doing that we want to outline two reasons for thinking that no interesting philosophical illumination is forthcoming from investigations into PRO (and related) constructions.

\textit{First Reason for Doubt: Back to “I”}

Here’s an initial reason for thinking that this data has no bearing on any interesting version of Essential Indexicality—or, more generally, on any interesting philosophical view: consider the standard replies to the question, “When can John be said to expect to win in a distinctly \textit{de se} way, rather than just to expect John to win?” Here are four ways in which e.g. Dilip Ninan (2010) characterizes the (allegedly) required \textit{de se} way of thinking about oneself (and there’s nothing idiosyncratic about Ninan’s characterization—he is just summarizing the literature):

- “\textit{De se} or ‘self-locating’ attitudes are thoughts one would characteristically express with a sentence containing the first-person pronoun I (me, my).” (p. 551)
- “They are thoughts about oneself when one thinks of oneself in the first-person way.” (p. 551)
- “...In order for (b) [John expects to win] to be true, John would need to think to himself, \textit{I am going to win the election}.” (p. 554)
- “In other words, (b) is true iff John has a \textit{de se} expectation to the effect that he will win.” (p. 554)
In some of these passages Ninan uses theoretical terms, such as “distinctly de se” and “self-locating.” These have no pre-theoretic meaning, they are terms of art introduced in the course of theorizing and for those (such as us) who are skeptical of the theories where these terms are used, they will be of little or no help. The same, we take it, goes for expressions like “thinks of himself in a first-person way.” For those who are puzzled by what this “first-person way” is (like us), that characterization alone won’t be helpful. There are, however, more substantive characterizations used by Ninan. We are told, for example, that de se thoughts are, “...thoughts one would characteristically express with a sentence containing the first-person pronoun I (me, my).” This strategy is illustrated by the claim that in order for “John expects to win” to be true, John has to think to himself I am going to win (as we have seen, Anand also appeals to dispositions to use the first-person pronoun to explain what he means by “de se” and “first person way of thinking”). We call this strategy the Back to “I”-characterization of PRO, or BTI. We have one simple, but we think important, observation about BTI: it makes discussion of PRO more or less irrelevant to the larger questions addressed in this book: i.e. whether there’s some kind of explanatory role for distinctly indexical content (and the de se in particular). BTI tells us that PRO has a content that’s distinctive in the relevant sense only insofar as utterances of sentences containing “I” have distinctive content of the relevant kind. But that’s the question that’s up for discussion. We claim they don’t, our opponents claim they do. BTI grants, in effect, that reflection on the semantics of PRO will do nothing to help adjudicate this disagreement.

In sum: when we are told that PRO functions in the relevant respects like “I,” the philosophical payoffs will come from reflections on the uses of “I”—there was no reason to take the detour of thinking about PRO. Seen in this light, it will seem that the cases outlined are simply illustrations of opacity. “Expects that” contexts are opaque. We know that as a result we can’t substitute, e.g. an occurrence of “I” for a co-referential expression. PRO functions like “I” in this respect. So understood, nothing new is learned from this data (other than that something “I”-like occurs in cases where you can’t see it (it’s unpronounced)).
Second Reason for Doubt: Lack of Connection to Philosophical Explanation

The second reason for doubting that the study of PRO constructions will be philosophically illuminating is that whatever interesting linguistic features the so-called “de se” readings have, they don’t connect in any interesting way to what philosophers have been interested in when they have been discussing the first person, indexicality, or perspectival thought. Two illustrations: first, there’s no interesting connection here to action, motivation for action, or rationalization of action. To see this, consider an example from Galen Strawson of creatures he calls “Weather Watchers” (1994, ch. 9). The key feature is that Weather Watchers have no capacity for action at all: they can’t form intentions or act on intentions. But they still have mental states: they represent the world, and maybe even desire the world to be certain ways. Note that Weather Watchers can be reported using PRO constructions. A Weather Watcher who sees a coming storm can expect to get wet. But that means that whatever the licensing kind of internal state is for the PRO construction, it’s not something that has to do with mental states that give rise to action.³

This point generalizes. The various features of “the de se” that are alleged to have wider philosophical implications, are not associated with PRO constructions. It is, for example, often assumed that distinctly “first-person thought” is associated with distinctive ways of acquiring self-knowledge. For reasons we go into in Chapter 7, we think this is wrong, but prior to that discussion it should be entirely obvious that the permissibility of PRO constructions is not tied in any way to special ways of learning about oneself. Imagine, for example, a being that doesn’t learn about itself via internal epistemic routes like proprioception and introspection, but learns everything about itself

³ It might be tempting to reply: No, the Weather Watcher’s states are the kind of states that produce action—they’re just not in a being that has any action capacities. But if, counterfactual, the Weather Watchers were to gain action abilities, then their de se states would be action-producing ones. Reply: The suggestion that in accepting the claim that the Weather Watcher expects to get wet, we’re making a judgment that this being would move around in certain ways were its physical capacities altered, is entirely unsupported by any evidence. We have no way of knowing what these creatures would do under those conditions, and there’s no reason to think that we make assumptions about that in asserting to the “expect to” claim. If it’s helpful, we can simply stipulate that Weather Watchers not only don’t act, but wouldn’t act, even if they gained capacities for action. All they care about is watching the weather.
by watching itself in a mirror—it is a creature that typically carries around mirrors for exactly this reason. Surely it, too, can expect to get wet, even though allegedly distinctive routes to self-knowledge are not available to it. In general, we venture the bold hypothesis that any philosophically substantive aspect of the de se can be severed from the PRO data in this way.

9.4. Brief Digression: Higginbotham on Self-Reflexivity

The current literature contains a number of interesting proposed semantics for PRO constructions. For most of these, the potential connection to philosophical issues of the kind we have been concerned with is brought out by informal glosses, such as BTI. It is, however, worth looking at one influential proposal by Jim Higginbotham. According to Higginbotham, the dividing line between PRO-permitting cases and non-PRO-permitting cases is self-reflexivity. He identifies distinctly de se states with those that are self-reflexive (and so PRO-permitting). The suggestion is that the semantic content of (a) is given informally by (b):

\[
\text{a) } A \text{ remembers } \phi \text{ing.}
\]

\[
\text{b) There is an event } e \text{ whose agent is } A \text{ and which is a remembering and whose content is that there is an event } d \text{ and the agent of } d \text{ is the agent of } e \text{ and which is a } \phi \text{ing.}
\]

In what follows, we will assume for the sake of argument that the semantic content of (a) is given by (b) and that this content is different from the content of “A remembers that A \( \varphi \text{ed} \)”, which isn’t self-reflexive. The point we want to make here is analogous to one made earlier: this distinction isn’t going to match up with the philosophically interesting de se phenomenon, because it doesn’t resolve the self-identity issue. To see why consider this case:

A is having an MRI, and can see his brain up on the screen (but doesn’t know it’s his brain). A sees a remembering state on the MRI screen, and thinks (and later remembers) “the agent of that state \( \varphi \text{ed} \)”. 
Note that Higginbotham’s proposed truth conditions for (a) are satisfied here: there is an event e whose agent is A and which is a remembering and whose content is that there is an event d and the agent of d is the agent of e and which is a φing. However, what is crucial for our purposes is this: this isn’t the kind of thing that anyone would think of as a distinctly de se state. The details here will be messy, but the underlying point is simple: Higginbotham tries to capture so-called “de se content” (distinctly first-person thought) by including a reference to the token mental state in the truth conditions of (a). But there’s no “de se story” about that reference—it’s just de re. But that means we can construct cases in which the subject is identity-confused about the token mental state, just like in standard Frege cases.4

9.5. Challenging the Data

In the remainder of this chapter we diverge somewhat from the overall strategy of this book. We take ourselves to have established that the issues about PRO won’t adjudicate any of the larger philosophical issues. That said, we suspect some might be attracted to the idea that even though it won’t settle any of those issues, it is an area in which very clean and clear data can be gathered that will then be useful for de se theorists at some point in theory-construction (and leave open just what point that is—the thought is: we have to start somewhere, and why not with something as clear-cut as PRO)?

Here’s a general philosophical strategy: there’s a supposed distinction in mental states that people are trying to characterize and understand. But it’s a complicated distinction, hooking on to many different areas of philosophy. So it would be nice to have a clean and simple tool for distinguishing between the “good cases” and the “bad cases.” At first glance, it can look like some simple linguistic constructions provide that tool. Perhaps we can distinguish the case in which Andy believes de se that his pants are on fire from the deviant case in which Andy sees his reflection in the mirror and thereby comes to believe,

4 An observant reader will note that this point is related to one we make about Higginbotham’s view in connection with immunity in Ch. 8.
without self-identification, that his pants are on fire by distinguishing the truth values of:

1) Andy believes his pants are on fire.
2) Andy believes Andy’s/that guy’s pants are on fire.

And in many contexts, the distinction between (1) and (2) will get people on to the distinction the de se theorist wants. But a little reflection on (1) and (2), and the broader behavior of attitude contexts, shows that this isn’t a clean and simple tool. There are many other contexts in which (1) and (2) are interchangeable. Figuring out what is pragmatics and what is semantics is, as the last century of philosophy of language attests, hard and messy.

So the de se theorist looks for other tools. Maybe it’s the use of the first-person pronoun that’s the clean and simple linguistic tool for separating good and bad cases? But again, careful examination of the phenomenon shows that the data is again messy and complicated, with controversy and unclarity about what is semantic and what is pragmatic. It’s at this point that the PRO constructions (and other similar phenomena) enter into the discussion. They represent the last best hope for a clean and simple linguistic tool that will allow separation of good and bad cases. In order to counter that admittedly vague thought or strategy, we end this chapter by showing that the PRO (etc.) data is much less clear-cut than is standardly assumed. It is data that has at least as much noise as any other linguistic construction containing indexicals.

9.6. Warm-Up: Familiar Frege Puzzles

Familiar Frege puzzles show that it is problematic to substitute co-referential singular terms in belief reports. It can be true to say that “Lois believes that Superman can fly” and false to say that “Lois believes that Clark Kent can fly.” Despite this, it is a familiar observation that belief reports are flexible and context-sensitive in such a way that there are some contexts in which such substitutions are acceptable. Here is the kind of case we have in mind:

Lois (who is unaware of the Superman/Clark Kent identity) says to Jimmy Olson (who is aware of the identity): “Clark Kent will be at a party at
Perry Whites house tonight.” Jimmy is then captured by Lex Luthor, and tortured to extract information on the whereabouts of Superman, so that Lex can plant some Kryptonite in advance. Finally Jimmy breaks, and tells Lex, “Lois said that Superman would be at the party at Perry White’s house.”

We tend to judge Jimmy’s claim as true, even though Lois used the name “Clark Kent,” and would not assent to the corresponding claim using the name “Superman.” Roughly, that’s because in the context, we’re not particularly concerned with tracking the details of Lois’s mental representations and their role in predicting and explaining her action. Rather, we’re concerned with tracking Lex Luthor’s actions, and the “Superman”-containing report allows us to do that.

9.7. “Expects To” Without the De Se

With these sorts of cases in mind, consider a variant on the expectation cases with which we began.

John is running for mayor of the local town. Earlier in the day, he gave a campaign speech, and he is now relaxing in the local pub. In an attempt to put the stresses of the campaign behind him, he has been drinking steadily all evening. The local news is showing on a television behind the bar, and is playing excerpts from John’s speech. The speech is a complete disaster—full of objectionable policies and errors on matters of facts. It’s obvious to us, sitting in the bar, that the speech has destroyed any chance John had of being elected. John, however, in his current inebriated state, finds the speech delightful. Too drunk to recognize himself on screen, he gestures broadly toward the television and declares, “That guy’s going to be our next mayor!”.

I turn to you and say, “Can you believe it? John’s so drunk he actually expects to win the election.”

If you share the view that the final report is true, then it is not a condition on its truth that John doesn’t make a mistake about who he expects will win. In this case, he says to himself: I will not win. So this is bad news for the proposal that expectation reports containing PRO can be true just in case the subject has a de se attitude. The relevant feature of this case is that the audience members, and the reporter in
particular, doesn’t care about how John thinks of himself; it’s a reporting context in which modes of presentation don’t matter.

9.8. “Admits To” Without the De Se

Consider the following case:

Clark Kent has lost track of the fact that he is Superman. In fact, he has been tracking Superman’s past activities, finding him a suspicious character. Superman comes to be on trial for sabotaging a power plant. Clark Kent is put on the stand (in front of a prosecutor, judge, and audience, all of whom are aware of his identity with Superman) and asked about what he’s discovered in his Superman-tracking. After Kent gives out one particular piece of information, very damning in context, the prosecutor declares, “Mr. Kent, don’t you realize that you’ve just admitted sabotaging the power plant?”

Again, there’s a mistaken identity: Clark Kent has lost track of the fact that he is Superman. He would deny the relevant “first-person thought” to himself, but “Kent admitted sabotaging the power plant” is true. Again, we find that the relevant kind of sentence containing a PRO construction is true, without an attitude anyone would call “de se” in the subject.

9.9. “Remembers” + Gerund Without the De Se

Consider the following case:

As a young child, John participated in and won a race at school. Many years later, John occurrently undergoes episodic memories from the race. The memories consist in flashes of visual imagery including the racetrack, the finish line, and a tape across the finish line breaking and falling to the ground, and in flashes of auditory imagery including cheers and someone saying, “Congratulations, you’ve won the race”. But John is uncertain who was being congratulated.
We think the case can be correctly described as follows: John remembers winning the race. If this is right, the PRO construction doesn’t require any kind of thought that’s plausibly characterized as distinctively first-person.5

9.10. “Only”–Binding Without the De Se

Consider the following variation on the previous case involving “only”–binding:

John and Frank are both running for mayor of Springfield, and both expect to win the election. They have also each been shown anonymized polling data regarding their campaigns, but are told that it’s data regarding two (fictional) candidates Bill and Edward in a different election. John studies (what he is told is Bill’s, but is in fact) his polling data, and concludes that Bill is likely to win. Frank studies (what he is told is Edward’s, but is in fact) his polling data, and concludes that Edward is unlikely to win. John and Frank are then given the opportunity to place bets on the outcome of the (fictional) election. John places a large bet on Bill, but Frank refrains from betting. When we ask why John bet and Frank did not, we are told, “Only John expected himself to win.”

We take it that this reply is correct, but, again, it doesn’t require the relevant kind of first-person thought.

9.11. Dream Reports Without the De Se

We’re not going to try to engage with all of the details of Anand’s account of de se pronouns in dream-report contexts, which include an exploration of the relations between English dream reports, logophors in Yoruba, and long-distance anaphors in Mandarin. We only want to make the observation that the English-language data is messy and

5 An anonymous referee for OUP suggested that “the predicate ‘winning the race’ might be being read de re. It’s saying that there is an event e such that e is a winning of the race, and John remembers participating in e (or something like that).” This is an interesting thought, but note that if this is the strategy appealed to here, it can just as easily be applied to the original cases from Higginbotham.
difficult, and not at all obviously in accordance with the generalization that Anand puts forth. Actually, Anand already more or less recognizes this in his work. Anand mentions cases such as:

I dreamed I was a carrot, and I was chopping me up for dinner.

I dreamed I was a roaring fire, and I sat down to watch my flames play in the darkness.

BLOCK requires that in these constructions, the third occurrence of “I” be interpreted as I-as-carrot, or I-as-roaring-fire. However, the semantic content of the sentence then makes the readings with these referential assignments rather odd. Anand reports that 60 percent of his data pool found such constructions “simply ungrammatical” or “pragmatically odd,” while 40 percent found them interpretable with the third “I” being either I-as-inanimate-object or I-as-person. The “simply ungrammatical” reaction is puzzling: one would think that at worst one would hear the claims as reporting very peculiar dreams in which a carrot-me chops up a person-me, or in which a surprisingly-animate fire-me watches the immolation of the person-me.

Our own informal surveys show a considerable range of reaction to these cases. The most common reaction was the recognition of the full range of ambiguities. Some speakers focused immediately on the carrot-chopping-up-person type readings, while others focused on the person-chopping-up-carrot reading. A substantial portion also got readings on which a carrot-me engages in an act of self-chopping. At least one speaker reported a global inability to read dream reports with multiple first-person pronouns so as to allow shift in referred dream character (and a preference for “myself” over “me” in such constructions).

So already:

• We don’t see much systematicity in the data, which to us suggests a pragmatic rather than a semantic phenomenon.
• We’re not sure what to make of a semantic generalization regarding the impact of de se phenomena, when that generalization only holds for (in the best case) 40 percent of the population.

We’re now going to present three more dream-report cases, and suggest that the net effect of these cases is to strongly suggest that there’s
no robust semantic phenomenon here, but just a cluster of pragmatic effects controlling pronoun interpretation in dream contexts.

Case 1: Brigitte Bardot Cases With Strong Contextual Effects

Director James Cameron says:

(REPORT) Last night, I dreamed I was Uma Thurman. I cast me in the lead role in a romantic comedy, opposite Hugh Grant.

Consider the following two dreams:

(DREAM₁) In his dream, James Cameron’s consciousness inhabits the body of Uma Thurman. Cameron-in-Thurman’s-body is then cast, with Hugh Grant, in a romantic comedy, with the casting done by James-Cameron’s-body.

(DREAM₂) In his dream, James Cameron’s consciousness inhabits the body of Uma Thurman. Cameron-in-Thurman’s-body then casts James-Cameron’s-body, with Hugh Grant, in a romantic comedy.

REPORT is a minor variant on a Brigitte Bardot sentence, and BLOCK should again predict that the only available reading is one true of DREAM₂, but not of DREAM₁. However, it is pragmatically distinguished from a Brigitte Bardot sentence by a context that makes DREAM₁ a more sensible reading, since on DREAM₁ the casting is being done by the director-in-the-dream, and the casting is done of the actor-in-the-dream.

We surveyed a number of speakers, asking them to determine whether REPORT was true of just one of DREAM₁ and DREAM₂, or of both, or neither. Almost none of our respondents took REPORT to be true only of DREAM₂, as predicted by BLOCK. In fact, respondents were more or less evenly split between those who found REPORT true only of DREAM₁, and those who found REPORT true of both dreams. (There was in addition a small minority who

---

6 We also checked for judgments on truth-but-misleadingness of REPORT regarding DREAM₁ and DREAM₂, and for which dream would be most naturally associated with REPORT.
found all such cases defective in some way or other.) Persistently in all our cases we found that about 30–40 percent of respondents took an extremely liberal position on which more or less any combination of pronominal reference was allowed.

Our conclusion from Case 1 is thus that selection of dream-character as referent for a first-person pronoun in a dream report is being driven primarily by pragmatic effects (the attempt to give a sensible interpretation to the dream), rather than any semantic constraint arising from interaction with the de se.

Case 2: Nominative and Accusative

Percus and Sauerland observe that the (putative) de se requirement on “I” in Brigitte Bardot sentences is lost when the first-person pronoun is moved inside an NP, as in: “I dreamed I was Brigitte Bardot, and my mother kissed me.” in which “my” reportedly can be interpreted either as “my-as-Brigitte-Bardot” or “my-as-dreamer.” Since “my” in this sentence does not c-command “me,” this is in keeping with BLOCK.

But the case also raises the possibility that the referential effects are being generated by a pragmatic “tracking” effect generated by the repeated use of “I.” To test this, we tried minor variants of Brigitte Bardot sentences in which the initial characterization of the dream uses the first-person pronouns in different cases:

(REPORT1) Last night I dreamed I was Brigitte Bardot, and I kissed me.

(REPORT2) Last night I dreamed Brigitte Bardot was me, and I kissed me.

(REPORT3) Last night I dreamed that Brigitte Bardot switched bodies with me, and I kissed me.

The resulting data was mixed and murky. In fact, most of our respondents found that in all three reports, both pronouns in “I kissed me” were referentially ambiguous, and there was relatively little shift in the distribution of judgments as we altered the form of the dream description. To the extent that there was a shift (in about 20 percent of the respondents), it was one that favored the kisser being the dreamer’s body...
in REPORT2 and REPORT3. Each of REPORT2 and REPORT3 were also found confusing by some respondents. Case 2 thus further suggests to us that the phenomenon here is primarily pragmatic. There is at least some suggestion that continuities in the morphological form of the pronoun are driving the interpretive procedure.

9.12. A Sketch of a Pragmatic Account

Here is a way to think of the structure of our opponents’ argument: a linguistic phenomenon is adduced, in which it is claimed that a certain construction is true of some scenarios and false of other scenarios. It is claimed that the distinction between the two classes of scenarios is the presence/absence of a distinctly de se mental state. It is then concluded that the linguistic phenomenon is sensitive to the presence/absence of such de se mental states. Our opponents then say: de se mental states are just whatever mental states are needed to make available “expects to” etc. constructions. The counter-data just given provides positive reason to doubt that this argumentative strategy will work. Our data shows that uses of PRO (and related) constructions do not require “de se readings” for their truth. The data is at least as messy as in other linguistic domains and a great deal of work is needed to distinguish semantic from pragmatic effects—this area provides no shortcut to the “good cases.”

It would go far beyond the scope of this work to provide a full theory of how this data should be properly structured, but we’ll end this chapter with a brief gesture towards what we suspect is the right kind of solution. Here’s a toy hypothesis:

(*)PRO constructions are used when there’s no funny stuff going on.

The thought is something like this: we have two constructions for reporting on A’s expectations regarding himself:

A expects A to win.
A expects to win.

7 At least one respondent reported that in REPORT2 it was unclear whether the dreamer was in the dream.
The first of these constructions is, for whatever reason, more “marked.” It’s longer, less usual, etc. So the thought is: we use the marked construction when there’s something funny going on. But then by implication, when we use the unmarked construction, we indicate that there isn’t anything funny going on. Identity confusion is a funny thing to have going on, so when there’s identity confusion, we tend to avoid the unmarked PRO construction. So suppose John is optimistic about his chances in the election, but then sees himself on television, fails to recognize himself as himself, and thinks that the guy on television is going to lose. (*) predicts, correctly, that (**) should be marked:

(**) John expects to win.

According to (*), we should reject (**) because there’s funny stuff (namely, identity confusion) going on. Note that there are standard (non-indexical-involving) Frege cases that show something like the (*) effects. So consider claims like:

a) When Superman gets in a fight, John expects Superman to win.

b) When Superman gets in a fight, John expects him to win.

c) When Superman gets in a fight, John expects a victory.

Now take a case in which John is watching fights in which Superman is sometimes entering, but in Clark Kent guise, and in which John is unaware of the Clark Kent = Superman identity. There’s some tendency to find (a) more acceptable than (c): somehow the identity confusion interferes more with the more condensed form than with the form in which there’s explicit use of the name.

Note that we can get the “only” binding cases with standard Frege cases as well. So suppose $A$, $B$, and $C$ have been captured by Lex Luthor. $A$ is aware of the Superman = Clark Kent identity, but $B$ and $C$ aren’t. $B$ and $C$ both think Clark Kent can’t save them, but that Superman can. $A$ thinks that Superman (whom he knows to be identical to Clark Kent) can’t save them, because he knows that Lex Luthor has kryptonite with him. Now consider:

Only $A$ thinks that Superman can’t save them.

There is a good reading of that sentence (at least, as good as the “only”—binding cases on PRO). So the more general phenomenon
may just be that “only” can bind some prominent reading. That non-identity-confused is a prominent reading is then not surprising.

There is obviously considerably more work to be done to show that this kind of pragmatic strategy will work across the board. Developing a full detailed theory goes beyond the scope of this book. And it is important to note that our two central points in this chapter stand independent of the viability of our toy theory:

1) We have no reason to think that theories of PRO will provide insight into the kinds of philosophical issues that have made Essential Indexicality seem important throughout philosophy.
2) The data is considerably more messy than standardly assumed.

Now, for those interested in accounting for that very messy data set, we recommend exploring a pragmatic strategy along the lines of our toy theory outlined here.

9.13. Brief Summary of This Chapter

This chapter has three central take-home messages (none of which we take to be very surprising or controversial):

1) No independent evidence for Essential Indexicality or the existence of irreducibly de se attitudes comes from the recent literature on the semantics of PRO constructions. That simply isn’t the goal of that literature. So if you agree with previous chapters of this book, the data about PRO constructions will provide no new philosophical insights about perspective, indexicality, or the first person.
2) We then noted, more or less as an afterthought, that the data is far from being as clean as we are sometimes led to believe. The

---

9. Here are some of the many questions to work on for someone interested in developing the pragmatic proposal: Is the “no-funny business” clause cancellable (for example, by using modifiers like “strictly speaking”)? Does it force us to be revisionary about a huge class of judgments and so end up being less elegant than alternative theories? Is there anything like a standard Gricean mechanism for the production of the proposed pragmatic effects? Thanks to Seth Yalcin for pointing out these and other challenges for the little toy-theory.
use of PRO constructions don’t map up with uses of, e.g. “I” in any clear and systematic way. The connections are messy and filled with pragmatic noise. In that respect it patterns more or less exactly with the data we have about substitutivity of singular terms in attitude contexts.

3) We sketched, but didn’t endorse, the beginning of a pragmatic strategy for cleaning up the data. A serious attempt of systematization of all the data would go very far beyond the scope of this book.
Chapter 10

The View from Everywhere

Our argument throughout this book is that the considerations that have come out of the Perry-Lewis tradition haven’t given us good reason to think that indexicality or perspectivality is a deep and philosophically interesting feature of contents\(^1\)—either of the information represented in contents, or in the way in which that information is stored in and is used by us. That’s a defensive move—no good reason has been given, but that’s not yet to say that the claim itself is false, or that there isn’t a good reason to be found elsewhere. The defensive move is our primary goal—we think the Perry-Lewis tradition has been influential across a wide range of philosophy in disproportion to the actual power of the considerations.

But we also think the substantive conclusion of the tradition is wrong, and that there is no deep or philosophically interesting notion of perspectival content. We think that contents are, and are used as, tools for representing (and, of course, sometimes misrepresenting) the objective state of the world. Some of the states represented are “perspectival” in the minimal sense that they are facts about our immediate environment, or facts about how things are in relation to us. Some of our representational systems are indexical in the minimal (Kaplanian) sense that they represent as they do in part in virtue of where they are situated in the world. But there’s nothing more to the phenomenon than that—fundamentally, all information is objective information, and is used indifferently by us as such.

What sort of contents should we expect there to be? That’s a question that can be approached by asking what we do with contents. An

\(^1\) Throughout this chapter we use “content” in an entirely non-technical, pre-theoretic way.
adequate account of the nature of content needs to make sense of the kinds of things we know, believe, hope, and want; of the many different ways in which we come to acquire and share information; of the way in which our beliefs interact with one another as our knowledge expands; and of the ways in which representing the world allows us to pursue our various aims and goals through action. At every turn, we think the best sense is made via purely objective, non-perspectival contents.

Our knowledge of the world is scissorshot. We don’t just know one kind of thing (spatial arrangements of objects, say) about one region of space (our immediate environments, say). If we did, it might make sense to use a representational system that used a specific kind of representation dedicated to that kind of knowledge (a “map of egocentric space”). But what we know, and the information we are constantly acquiring, are little, often disunified, pieces from all over.

We know something about the spatial relations between London and Oslo—that Oslo is north and east of London. And we know something about the spatial relations between London and Juneau—that Juneau is north and west of London. But we may well be partially ignorant of the spatial relations between Oslo and Juneau, knowing that Oslo is east of Juneau, but uncertain whether it is north or south of Juneau. And, of course, we might have no idea where any of the three cities are relative to us.

Fitting these things we know into some sort of dedicated special-purpose perspectival representation system won’t be easy. We can’t know about the spatial relations among London, Oslo, and Juneau by placing them on a map of egocentric space, if we don’t know where they are relative to us, or on a map at all, if we don’t know how to place Oslo and Juneau relative to one another. But fitting these things into a generic all-purpose representation system, which tracks information in the form of objective truth-conditional propositions, isn’t a problem. It’s easy to represent that Oslo and Juneau are both north of London, without representing the north–south relations between Oslo and Juneau, if we do so by endorsing the propositions that Oslo is north of London and that Juneau is north of London while refraining from endorsing either that Oslo is north of Juneau or that Juneau is north of Oslo.
This is a ubiquitous feature of our knowledge. We know that the car is old without knowing whether it is a convertible; we know that the sign is red without knowing what shade of red; we know that Napoleon ought to have managed the invasion of Russia better without knowing how he did manage it. What we know fits no reliable pattern, and requires a representational system that doesn’t require a pattern, and doesn’t ask that the things we know be specially related to us in some way. It’s a big world—we don’t have a labeled map of it, but just scattered nuggets of information about it.

We are omnivorous and indifferent consumers of information. There is a tendency—raised to the status of high theory in functionalist accounts—to focus on the role of information in our cognitive lives as a tool for realizing ends through action. That’s certainly one important aspect of cognitive economy of information, but it’s far from the only role. We don’t collect information just with an eye toward achieving (short-term or long-term) goals—we collect information because we can’t help it. That’s just the kind of beings we are: scavengers for data. We absorb information wherever we find it, without concern for how that information stands in relation to us or to other pieces of information. We may sometimes do some systematization of that information, but such organizations are local, transitory, and hard work—not the ordinary function of our representational system. What is constant is that there is a world of objects and relations, and we accumulate bits and pieces of the goings-on of that world.

Our scattershot knowledge of the world gives us need for a system of representation that deploys non-perspectival information. But maybe that’s just to say that the perspectival information is one special part of what we know—that our knowledge mingles perspectival and non-perspectival information. That’s incompatible with globalist pictures of perspectivity, on which indexical content is a ubiquitous feature of our knowledge, but compatible with more localist pictures. However, at the same time that our knowledge is scattered in content and topic, it is fully integrated in its availability. Anything that we know can be brought to bear on anything else that we know. Our knowledge of the spatial relation between London and Juneau may be partial, and an isolated nugget of spatial information. But no matter how partial and isolated, it is available for inferential integration with other things.
we know. If we know that everything north of London is cold, we’re in a position to go on to know that Juneau is cold.

If our knowledge is a mixture of perspectival and non-perspectival information, then there’s an extra challenge to be overcome in getting everything potentially inferentially related, especially if perspectival information is *deeply and importantly different* from non-perspectival information. But if our knowledge consists solely of pieces of objective information (some of it about us and our local environment, some of it not), then the inferential integration challenge is easily met.

Some of the information integration we do is *interpersonal*. We’re not just informational scavengers, we’re *social* informational scavengers, passing what we know back and forth among ourselves. That’s bad news if some of our contents have some proprietary epistemic or informational architecture, marked as representing the world as it is for one of us, or as obtained in some particular way. (It’s no accident that, for example, centered-worlds stories of content have special challenges in providing an adequate story about communication.) What we want is fully portable information, and information that coordinates on the world, rather than on us, is best suited for that task. It’s also bad news if some of our contents are essentially tied to particular contexts. We want to pass information around widely across time and space—if the information is objective, and can be retrieved and accessed indifferent-ly by everyone, that is possible, but if the information needs to be assessed relative to some ur-context, later recipients won’t know what to do with it.

But even *intrapersonally* it would be an inconvenience to mix perspectival and non-perspectival contents. If perspectival contents are *informationally* marked, by being in the business of representing not *how things are*, but rather *how things are from here* (that is, if they’re marked by having relativist truth conditions), then we need special inferential hygenics to make sure that absolutist and relativist truth conditions play nicely together. If perspectival contents are *epistemically* marked, by being tagged as contents derived from certain proprietary personal informational channels, then we need special epistemic hygenics, to make sure that we keep proper track of what data came from what source. But we’re not careful in either of these ways. We don’t track what we learned where or how, and we freely combine all of what
we know using global inferential procedures. Information from our earlier selves isn’t interestingly different from information from other people—in both cases, objectivity ensures portability.

One (but only one) of the important ways in which we gain information is via perception, and it’s a natural and common thought that perception will be a special source of perspectivality, because perception reveals how the world is *around here*, seen in *this way*. But this natural thought not only can but should be resisted. Perception is informationally richer with the local than with the distal. (Up to a point: get too close to some of the sensory organs, and information degrades. And to a certain degree: perceptual information quite regularly and robustly comes from an object some 93 million miles away.) But *where* the perceptually represented items are (with respect to the perceiver) is a poor guide to *how* they are represented. Some relational features do make their way into perceptual content, but for the most part perception presents the environment objectively—what we receive through vision is information about the features of, and spatial relations among, various objects (among them, us, and parts of our body). What we primarily perceive is, for example, chairs and tables in certain relations to one another—this is why we can, on the basis of that perception, navigate the chairs and tables from an angle different from the original angle of perception, or draw a “god’s-eye” map of the chairs and tables. Surely this is what we would expect and need of a perceptual system that would help us navigate a world on which we possess no fixed perspective.

Secondarily, of course, perception reveals relational information between the environment and the perceiver. The tree is perceptually represented as taller than the lamp, but the lamp also in some sense looks “taller than the tree,” by way of visually disclosing a complicated combination-of-size-plus-proximity-to-perceiver feature. But this relational information is clearly a secondary aspect of perception. It is hard to extract—one reason drawing is hard is because our perceptual systems so dominantly give us the objective features of things that one must focus and train to recover the more relational features, reproduction of which is necessary for a convincing drawing (the tree is seen as taller, but the drawing of the tree must be shorter than the drawing of the lamp).
If perception is to be of any use to beings like us, perceptual contents need to transition into perceptual memories and perception-based beliefs. But perceptual memories retain objective relations much more robustly than perspectival features. When you recall the look of the inside of your house, you are likely to be presented with quasi-visual impressions of spatial arrangements and appearances of objects from no certain or stable perspective—where things are relative to one another is what persists in memory, not where they were relative to you when seen. And if perception is to integrate with belief, then all of the earlier reasons that belief traffics in objective contents are also reasons why a helpful perceptual content also traffics in objective contents.

The contents of our mental states, as already noted, should not come epistemically inflected. But furthermore, the epistemic relations among those states should not be perspectivally inflected. Evidential relations are no respecters of person. Whatever it is that makes one thing evidence for another, it is not that both things are mine (or that one is yours and the other is mine). The evidential relations reside in the pieces of evidence, not in the ownership relation between person and evidence. If evidence is wholly worldly (external facts, or mental states representing external facts), then the evidential relation can be, or be a reflection of, the relation of necessary truth preservation. If evidence is deeply experiential, then empiricism is achieved by having the evidential relation carried by, for example, the phenomenology of the experience. If evidence is holistic, then a Bayesian picture can be achieved by having the evidential relation hold between total credence states. But any of these epistemic pictures is made a mockery of if first-person states are given epistemological center stage, because then the holder of the evidence mysteriously enters into the evidential relation.

Your beliefs and experiences are evidence for you and not for someone else, not because their being yours gives them an evidential status, but just because, as a contingency of our causal setup, they are available for you, and not for someone else. If we were causally organized so that your neighbor could form beliefs (in the right way) on the basis of your perceptual experience, then those beliefs would be evidentially supported by your experience, across the person boundary. It’s this impersonality of the evidential relation that makes possible reductive memory and evidence based accounts of personal identity—if
The View from Everywhere

evidential relations were intrinsically personal and person-bound, then we would need a (mysterious) prior notion of what a person was. The impersonality of the evidential relation also means that when the causal setup is right, your states can be evidence for your neighbor’s beliefs—testimony, for example, is merely causally indirect perception. We are mere conduits for evidence—we do not so much infer, as we are inferred through.

We represent the world not just because we are actors in the world, but our roles in the worldly cast do impose one important success condition on our representations. And that imposition is again objective, not perspectival and personal. To act, we need proper engagement between our representations of how the world is and of how it ought to be. But how the world ought to be—the ends that our actions subserve—is not essentially a personal matter. Many of our goals are indeed local—we typically have better information about the goings-on of our local environment, and hence better awareness of the deficiencies of that environment; and we are as a matter of biological contingency typically more moved by awareness of local deficiencies than by awareness of more distant deficiencies.

But this locality should not be exaggerated. You may care more about your own pain than the pain of others—undergoing a pain is a quite effective causal mechanism for creating anti-pain motivations. But you also care greatly about the pain of your loved ones. And not because they are your loved ones—a motivational path of the form “So-and-so is in pain; so-and-so is importantly related to me; so I should alleviate so-and-so’s pain” is psychopathically narcissistic, not normal. The loving relation between you and your loved ones is a mechanism for creating sensitivity to the motivational power of their suffering, not the source of that motivational power. It is a mistake to think that all interest is (overtly or covertly) self-interest. That’s manifestly not how it is, and it’s not how you would expect it to be for beings who came about the way we did, as products of causal forces shaped by preservation of genetic resources rather than individuals, and as members of species whose survival methods were largely communal rather than individual.

If only special first-person states could give rise to action in us, we would be strangely vulnerable to the environment. Threats and
opportunities arise and change unexpectedly and regularly, and we need to stand globally ready to react and exploit. But this means that everything we know about the world needs to be potentially motivational, or available for interaction with our motivations. Creatures that run away from predators only when they form special de se predator-avoiding states are likely to end up eaten.

In short: our view on the world is not primarily a view from a perspective. Our beliefs and desires are not organized around us. They are instead organized around the world itself (which is just a way of saying that they are not by their nature organized at all). Our view is a view from everywhere. Our nature is not deeply advantageous on the world, but as one among many occupants of the world. We see ourselves along with everything else. Some things are seen more clearly and some less. The unclarities in our views of ourselves are at least as prominent as the clarities.

So then, what of perspective and indexicality? These are real phenomena, they just aren’t deep phenomena. Some of what we know about the world is perspectival, in that we ourselves are hidden relata in it. When you know that the sound is to your left, what you know is perspectival, because it is knowledge of a spatial relation between you and the sound, not of a monadic property of the sound. But the specifically relational nature of the representation is hidden from you. It is the hiddenness that creates the perspectivity. Your knowledge that you are shorter than Mount Everest is also knowledge of a relation between you and something else, but it does not seem a perspectival bit of knowledge—we know just as much as you, and in the same way as you, that you are shorter than Mount Everest.

So the interest here is in the phenomenon of hiddenness. But hiddenness of this sort is a ubiquitous and expected part of our representational lives. The child can know that Joe is his uncle while the relational structure of unclehood remains hidden from him. The child can think about water while the microstructure of water remains hidden from him, and thus fail to connect his water thoughts with his H₂O thoughts. We can see Nora, and see Nora again, but have the sameness of subject of those two seeings hidden from us, and fail to recognize her one or both times.

In representing things in the world, we place ourselves in cognitive contact with those things, and create certain capacities to recognize
and manipulate those things. But the cognitive contact and the concomitant capacities are imperfect and subject to occasional failure. There can be a tendency to try to explain the failure of one skill via the successful application of another skill—we failed to jump over the high bar because our successful \textit{trying to jump over} was for extrinsic reasons not identical to a \textit{jumping over}; we failed to recognize Clark Kent as Superman because our successful \textit{recognizing of the Clark Kent and Superman guises} was for extrinsic reasons not identical to a \textit{recognition of Clark Kent}. But this is a misguided and hopeless picture of skills. Our failures are not always underwritten by prior successes — quite often, they are simply and irreducibly failures. Success stands just as much in need of explanation as failure.

Perspectivality, then, is just our occasional failure to see everything about the world there is to be seen. The seeming \textit{nearness}, rather than \textit{near-to-me-ness}, of the tree is of a piece with the failure to see the backs of objects—a manifestation of our regrettable, but hardly crippling, epistemic limitations. A full understanding of our epistemic limitations is a deep and interesting topic, but the perspectivality per se is not. Indexicality is, if anything, even less deep than perspectivality. Representational states represent by virtue of various features of those states. The sentence “Snow is white” represents what it does by virtue of the words it is composed of, the meanings of those words, the syntactic pattern in which those words are arranged, and the semantic effect of that syntactic pattern. (All of these facts may in turn hold by virtue of yet more facts.) Some features of “Snow is white” are representationally significant, and others are not. (The particular choice of font, for example, is not representationally significant.)

So it could be that the representationally significant features of a particular representational state are only those that it shares with other representational states of the same kind, for some salient kind. That’s how it is with “Snow is white.” That particular token sentence represents as it does entirely by virtue of features that it shares with other tokens of “Snow is white,” such as that one. But it needn’t always be that way. Particular representational states have many features they don’t share with their kind-cousins (and, as we stressed in Chapter 2, for non-linguistic cases even finding a relevant kind can be difficult). So for some sentence types, such as “I am hungry,” the representational
content of tokenings will depend on features of the tokens that are not shared with other tokens of the same type.

Of course, representationally significant features of a representational state must be features of that state. That’s a triviality, but it’s in danger of looking deep. It can lead one to think that somehow locality matters—that representations always represent by virtue of how things are with respect to them, rather than by virtue of how things are elsewhere. But that’s just because they are the representers, so they need to have the representation-securing features. When you add to this the fact that we ourselves are typically quite proximate to the representational states we produce, consume, and utilize, it can look as if we ourselves are somehow implicated in the representing. Add a bit of self-involvement to the mix, and suddenly indexicality can seem quite important. But our proximity to the representations we use is clearly a contingent, and indeed frequently defeated, fact—there is no deep or essential connection between indexicality and the self.

Perspectivality and indexicality are real but shallow. They are predictable consequences of imperfect agents going about their task of trying to represent the world as it objectively and impersonally is. They are phenomena that not only make sense, but make the best sense, when placed within a general picture of content as objective representation of the world.

The End.


Bibliography


—— (2009), “*De Re* and *De Se*,” *Dialectica* 63: 249–269.


INDEX

action, philosophy of 11, 14
action explanation 3, 27, 31, 33, 36 n. 5, 39, 41, 42 n.10, 46, 48, 50, 55, 56 n. 20, 60, 91
impersonal 35–7, 39–40, 46, 55
personal 36, 46
Action Inventory Model 49–52
action rationalization, see action explanation
agency 7–8, 12, 14, 18, 26, 30–57, 109
Akerman, J. 17
Anand, P. 152 n. 1, 154, 156, 158, 165–6
Argument from Phenomenal Revelation, see perceptual content
Argument from Weak Revelation see perceptual content
Arntzenius, F. 13
attitudes 18, 23, 41, 51, 60, 65–7, 70, 86–7, 90–1, 100, 111, 121–2, 162, 171–2
de dicto 90
de re 90, 105
de se 4–6, 13, 19–20, 91, 105, 152–4, 157, 163–4
de tu 104 n. 7, 105
audience 24, 104–5, 120

belief 24, 32–3, 39, 48–55, 60, 87 n. 2, 106, 136–8, 178
de dicto 96
demonstrative 134
de re 99
de se 4, 62 n. 1, 96, 140
indexical 11, 23, 40, 46, 133
Block, N. 71
Braun, D. 70, 76–7
Brogaard, B. 11 n. 5, 12, 146–50

Campbell, J. 131, 134
Caplan, B. 149 n. 6
Cappelen, H. 93, 144 n. 3
Castañeda, H. 3–4, 8–9, 11, 17, 39, 86, 152
causation 34–5, 110, 118, 132, 135, 178–9
centered worlds 11–12, 86, 92, 97–8, 102–8, 114, 116–25, 135, 137, 148–50, 176
Chalmers, D. 7, 12, 71, 146–50
character, Kaplanian 16
Chierchia, G. 152–3
Christofidou, A. 13
Cohen, S. 17
concepts 7, 16, 21, 30, 45, 73 n. 8, 74–5, 91, 134–5, 145
consciousness 7–8, 12, 26
content, Kaplanian 15
corpusensitivity 16–17, 21, 23, 24 n. 3, 27–8, 35 n. 3, 42, 81, 162
contextualism 11, 92, 144 n. 3, 150
co-reference 14 n. 8, 18, 28, 33, 39, 58–60, 67 n. 6, 77, 84, 101, 158

Davidson, D. 70
de se
and agency 30–57
attitudes 4–6, 13, 19–20, 91, 105, 152–4, 157, 163–4
beliefs 4, 62 n. 1, 96, 140
and centered worlds 86–108
implicit 137
irreducible 2–6, 19–20, 140–1, 146, 171
mental states 11, 35 n. 3, 120, 156, 169
and PRO construction 151–72
desire , 36 n. 5, 37 n. 6, 46–55, 107, 109–24, 180
demonstratives 10, 16, 26–7, 42, 73, 128, 134
“Department Meeting” case 65, 80
descriptivism 10, 71, 99
diagnostic distinctiveness 59, 69, 77, 81
Fregean 59, 69–73
dream reports 154, 156, 165–9
<table>
<thead>
<tr>
<th>Index</th>
<th>Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Egan, A.</td>
<td>11–12, 17, 65 n. 4, 92, 105–6, 111 n. 1, 115 n. 4, 148, 163</td>
</tr>
<tr>
<td>“Election” case</td>
<td>155–6, 163, 165</td>
</tr>
<tr>
<td>Elga, A.</td>
<td>13</td>
</tr>
<tr>
<td>epistemology</td>
<td>13–14, 19, 71, 84, 127–9, 140, 178</td>
</tr>
<tr>
<td>formal</td>
<td>13</td>
</tr>
<tr>
<td>Evans, G.</td>
<td>10, 24, 71, 73, 79, 128–32</td>
</tr>
<tr>
<td>Explanation, see action explanation</td>
<td></td>
</tr>
<tr>
<td>and consciousness</td>
<td>8</td>
</tr>
<tr>
<td>content</td>
<td>26, 50, 122, 132</td>
</tr>
<tr>
<td>knowledge</td>
<td>13</td>
</tr>
<tr>
<td>perspective</td>
<td>6</td>
</tr>
<tr>
<td>pronoun</td>
<td>23, 154–8, 162, 166–9</td>
</tr>
<tr>
<td>Fodor, J.</td>
<td>42, 53 n. 16, 71</td>
</tr>
<tr>
<td>Forbes, G.</td>
<td>10</td>
</tr>
<tr>
<td>Frege, G.</td>
<td>10, 18, 70–3, 78–81</td>
</tr>
<tr>
<td>Cicero / Tully</td>
<td>80</td>
</tr>
<tr>
<td>Hesperus / Phosphorus</td>
<td>58, 100, 120–2</td>
</tr>
<tr>
<td>Superman/Clark Kent</td>
<td>28, 33, 39, 46–9, 53 n. 16, 58, 60–8, 81, 100–1, 162–4, 170, 181</td>
</tr>
<tr>
<td>Fregean senses</td>
<td>17 n. 9, 71, 79, 112</td>
</tr>
<tr>
<td>descriptive</td>
<td>71</td>
</tr>
<tr>
<td>robust</td>
<td>71, 79</td>
</tr>
<tr>
<td>Fregeanism</td>
<td>53 n. 16, 58–85, 96</td>
</tr>
<tr>
<td>functionalism</td>
<td>19, 24, 53 n. 17, 91, 109–26</td>
</tr>
<tr>
<td>fancy</td>
<td>114–19</td>
</tr>
<tr>
<td>folk</td>
<td>111–14, 117</td>
</tr>
<tr>
<td>Fregean</td>
<td>120–1</td>
</tr>
<tr>
<td>rescue-by strategy “Defense”</td>
<td>110, 120, 125</td>
</tr>
<tr>
<td>Grim, P.</td>
<td>13</td>
</tr>
<tr>
<td>Hamilton, A.</td>
<td>13</td>
</tr>
<tr>
<td>Hare, C.</td>
<td>13</td>
</tr>
<tr>
<td>Hare, R. M.</td>
<td>11</td>
</tr>
<tr>
<td>Hawthorne, J.</td>
<td>17, 93, 144 n. 3</td>
</tr>
<tr>
<td>Heck, R.</td>
<td>79 n. 9</td>
</tr>
<tr>
<td>“Heimson” case</td>
<td>22–3, 64, 87 n. 2</td>
</tr>
<tr>
<td>Higginbotham, J.</td>
<td>11, 131, 135, 138–9, 155, 160–1, 165 n. 5</td>
</tr>
<tr>
<td>immunity to error through</td>
<td></td>
</tr>
<tr>
<td>misidentification</td>
<td>13, 15, 19, 82, 127–39</td>
</tr>
<tr>
<td>impersonal incompleteness claim (IIC)</td>
<td>37–41, 46–9, 56–7</td>
</tr>
<tr>
<td>theses weaker than</td>
<td>39–40</td>
</tr>
<tr>
<td>indexicality</td>
<td></td>
</tr>
<tr>
<td>shallow</td>
<td>1–2, 5, 83, 182</td>
</tr>
<tr>
<td>spatial</td>
<td>142–5</td>
</tr>
<tr>
<td>temporal</td>
<td>12, 27, 53–4, 65–7, 80</td>
</tr>
<tr>
<td>information gathering</td>
<td>15, 82, 86, 146</td>
</tr>
<tr>
<td>intentionality</td>
<td>115 n. 5</td>
</tr>
<tr>
<td>Johnston, M.</td>
<td>12</td>
</tr>
<tr>
<td>Kamp, H.</td>
<td>11</td>
</tr>
<tr>
<td>Kaplan, D.</td>
<td>10, 17, 26–7, 58, 77, 79</td>
</tr>
<tr>
<td>Kaplanianism</td>
<td>9–11, 15–16, 38, 81, 173</td>
</tr>
<tr>
<td>Katz, J.</td>
<td>73</td>
</tr>
<tr>
<td>Kenny, A.</td>
<td>11</td>
</tr>
<tr>
<td>knowledge</td>
<td>7, 30, 63, 67, 88–90, 96, 128–32, 174–6, 180</td>
</tr>
<tr>
<td>de dicto</td>
<td>4</td>
</tr>
<tr>
<td>de se</td>
<td>4</td>
</tr>
<tr>
<td>first-person</td>
<td>13</td>
</tr>
<tr>
<td>self-knowledge</td>
<td>84, 159–60</td>
</tr>
<tr>
<td>Kratzer, A.</td>
<td>17</td>
</tr>
<tr>
<td>Kretzmann, N.</td>
<td>13</td>
</tr>
<tr>
<td>Kripke, S.</td>
<td>10, 69, 71, 75–7, 79, 85</td>
</tr>
<tr>
<td>Epistemic Argument</td>
<td>75–7</td>
</tr>
<tr>
<td>Problem of Ignorance</td>
<td>75–6</td>
</tr>
<tr>
<td>Problem of Misdescription</td>
<td>75</td>
</tr>
<tr>
<td>Lakoff, G.</td>
<td>156</td>
</tr>
<tr>
<td>language of thought</td>
<td>24</td>
</tr>
<tr>
<td>language, philosophy of</td>
<td>9, 11, 21–7, 162</td>
</tr>
<tr>
<td>Larson, R.</td>
<td>70</td>
</tr>
<tr>
<td>Lee, G.</td>
<td>12</td>
</tr>
<tr>
<td>Lewis, D.</td>
<td></td>
</tr>
<tr>
<td>“Attitudes De Dicto and De Se”</td>
<td>3, 6, 19, 67, 86–7</td>
</tr>
<tr>
<td>On the Plurality of Worlds</td>
<td>95</td>
</tr>
<tr>
<td>“Reduction of Mind”</td>
<td>110</td>
</tr>
<tr>
<td>“Veridical Hallucination and Prosthetic Vision”</td>
<td>142 n. 1</td>
</tr>
<tr>
<td>“Lingens the Amnesiac” case</td>
<td>4, 62–4, 87, 89–91, 98, 141–2</td>
</tr>
<tr>
<td>Name</td>
<td>Refs</td>
</tr>
<tr>
<td>------------------</td>
<td>-------------------------------------------</td>
</tr>
<tr>
<td>Richard, M.</td>
<td>11 n. 5, 70</td>
</tr>
<tr>
<td>Rumfitt, I.</td>
<td>80</td>
</tr>
<tr>
<td>Russellianism</td>
<td>9, 15, 38, 70</td>
</tr>
<tr>
<td>Salmon, N.</td>
<td>70</td>
</tr>
<tr>
<td>Sauerland, U.</td>
<td>156, 168</td>
</tr>
<tr>
<td>Schiffer, S.</td>
<td>79</td>
</tr>
<tr>
<td>Schlenker, P.</td>
<td>17</td>
</tr>
<tr>
<td>Schroeder, T.</td>
<td>149 n. 6</td>
</tr>
<tr>
<td>The Selection Problem</td>
<td>49–51</td>
</tr>
<tr>
<td>self-ascription</td>
<td>10–12, 86–108, 120, 122, 124–5, 145</td>
</tr>
<tr>
<td>self-consciousness, see consciousness</td>
<td></td>
</tr>
<tr>
<td>self-location</td>
<td>2–4, 94, 96, 106, 157–8</td>
</tr>
<tr>
<td>self-reflexivity</td>
<td>11, 160–1</td>
</tr>
<tr>
<td>semantics</td>
<td>10–12, 60, 80 n. 10, 84, 94, 98, 151–72</td>
</tr>
<tr>
<td>metasemantics</td>
<td>110, 117–24, 138</td>
</tr>
<tr>
<td>Shoemaker, S.</td>
<td>6, 12, 127, 129</td>
</tr>
<tr>
<td>Sider, T.</td>
<td>66</td>
</tr>
<tr>
<td>Siegel, S.</td>
<td>8, 142–8</td>
</tr>
<tr>
<td>Sleeping Beauty problem</td>
<td>13</td>
</tr>
<tr>
<td>Soames, S.</td>
<td>70, 84</td>
</tr>
<tr>
<td>Smith, J.</td>
<td>13</td>
</tr>
<tr>
<td>speaker</td>
<td>16, 24, 41, 64–5</td>
</tr>
<tr>
<td>Stanley, J.</td>
<td>122</td>
</tr>
<tr>
<td>Stalnaker, R.</td>
<td>115</td>
</tr>
<tr>
<td>Strawson, G.</td>
<td>159</td>
</tr>
<tr>
<td>Strawson, PF.</td>
<td>129</td>
</tr>
<tr>
<td>substitution</td>
<td>salva veritate 28, 33, 39, 57–8, 83–4</td>
</tr>
<tr>
<td>super indexical</td>
<td>26–7</td>
</tr>
<tr>
<td>“Thank goodness that’s over” case</td>
<td>66–7</td>
</tr>
<tr>
<td>thinker</td>
<td>8, 24, 26–7, 38, 41, 64, 86</td>
</tr>
<tr>
<td>Titelbaum, M.</td>
<td>13</td>
</tr>
<tr>
<td>“Two gods” case</td>
<td>67–8, 87–106</td>
</tr>
<tr>
<td>Tye, M.</td>
<td>70, 146</td>
</tr>
<tr>
<td>van Fraassen, B.</td>
<td>6</td>
</tr>
<tr>
<td>Velleman, D.</td>
<td>8, 13, 31, 38</td>
</tr>
<tr>
<td>“Weather Watchers” case</td>
<td>159</td>
</tr>
<tr>
<td>Wittgenstein, L.</td>
<td>127–30</td>
</tr>
<tr>
<td>Yalcin, S.</td>
<td>12, 107, 122–3, 171 n. 8</td>
</tr>
</tbody>
</table>