

Ways of Thinking

An Essay on Referential Coordination

Henry Clarke
Department of Philosophy
UCL

Submitted for the degree of Doctor of Philosophy

Declaration

I, Henry Clarke, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Abstract

Referential coordination occurs when a thinker is rational in treating her thoughts as being about the same thing. This is manifested primarily in the thinker's dispositions to make inferences, paradigmatically the disposition to infer an existential generalisation conjoining two or more properties without recourse to an additional premise concerning an identity. It therefore presents an indispensable way for identity to figure in thought. This topic is often addressed in the form of discussions of so-called Frege cases, identity judgements, or coreference *de jure*. I argue that referential coordination should be treated as an independent and prior explanatory problem. The problem referential coordination presents is to explain the rationality of the paradigmatic inferential dispositions. I discuss three prominent theories of thought in relation to this problem: the appeal to propositional contents akin to Frege's notion of sense; the appeal to mental representations that can be typed in some way; and the appeal to mental files and their functional properties. Representatives of these theories fail to provide an explanation that is at once non-circular, psychologically realistic, and sufficiently general. I propose an alternative *coordination functions* explanation. This uses an amended version of mental file theory that distinguishes between mental files and file predications, and combines this with an apparatus of defaults and defeaters familiar from entitlement epistemology. File predications, the associations of files with bits of information, serve as the basis of the paradigmatic inferential dispositions, and so have normative functional properties that provide a default indication of sameness of reference open to defeat by conflicting information. This relatively deflationary explanation is distinctive in dispensing with any explanatory notion of a concept. It can be extended to providing a similarly deflationary account of the rational role of identity judgements and thoughts about oneself and one's immediate environment.

Acknowledgements

For financial assistance while this thesis was written, I am grateful to the Arts and Humanities Research Council for a Doctoral Studentship (2011-2013), and to the Royal Institute of Philosophy for a Jacobsen Studentship (2013-2014). I am also grateful to UCL for granting me additional time to complete the thesis during 2015-2016, and to Dr. Fiona Leigh for lending her time to assist in applying for it.

I would like to thank Prof. Christopher Peacocke, Dr. Ian Phillips, and particularly Dr. Daniel Rothschild, who supervised the main bulk of the work. Their comments and suggestions are woven into the thesis at every point, and it would not be what it is without their supervision.

I have gained immeasurably from the supportive and friendly environment of the Philosophy department at UCL. It is impossible to name each of my colleagues who aided the process in some way, but I would like to thank all those who attended Work in Progress and Departmental seminars at which I presented early drafts for their rigour and generosity in questioning. Discussions during meetings of reading groups on Peacocke's *Truly Understood* and Recanati's *Mental Files* gave rise to early forms of arguments that appear in this thesis. I would especially like to thank Dr. Ed Nettel for many stimulating conversations and much useful advice during our time together at the department.

Many thanks to Abigail Cooke for (voluntarily!) lending a hand with proof-reading.

Writing a doctoral thesis (perhaps especially in philosophy) is an occasionally isolating and often dispiriting task. This probably does not show up in the thesis, though it did show up a great deal in my life in the last four and a half years. I would like to take this opportunity to encourage anyone reading this to ask themselves how they are doing; if you need some help, do ask for it. You will be surprised, as I was, by how kind people can be.

My endless thanks to Lucy Lopez, who helped in ways too many to count, and to my parents, for everything.

Amsterdam, March 2016

Contents

Introduction	7
Chapter 1 - The Problem of Referential Coordination.....	15
1.1 Introduction	15
1.2 Referential Coordination as an Explanatory Problem	15
1.3 Related Problems	27
1.4 Questions and Theories	40
1.5 Assumptions, Requirements, and Desiderata	47
1.6 Conclusion	50
Chapter 2 - Concepts.....	51
2.1 Introduction	51
2.2 What Are Concepts?	51
2.3 Conceptual Capacities	55
2.4 Alternative Answers.....	59
2.5 Philosophical vs. Psychological Concept Theory.....	62
2.6 Eliminating Concepts	64
2.7 Conclusion	66
Chapter 3 - Sense Theory.....	67
3.1 Introduction	67
3.2 Developing Frege's Notion of Sense.....	68
3.3 The Handles Challenge	73
3.4 Theories of Grasp	77
3.5 Circularity Problems	82
3.6 Conclusion	96
Chapter 4 - Typed Representations Theory.....	97
4.1 Introduction	97
4.2 Typed Representations as an Alternative to Sense Theory	98
4.3 The Typing Challenge	101
4.4 The Syntactic Typology	103
4.5 The Originalist Typology.....	108
4.6 The Neo-Empiricist Typology	113
4.7 Sameness Markers	120
4.8 Conclusion	126

Chapter 5 - Mental File Theory.....	127
5.1 Introduction	127
5.2 Explanation by Mental File	128
5.3 Basic Model Explanations	136
5.4 The Indexical Model	140
5.5 The Generality Problem	149
5.6 Conclusion	157
Chapter 6 - Coordination Functions Theory	159
6.1 Introduction	159
6.2 Predications, Requirements, and Functions.....	160
6.3 Evaluation	178
6.4 The Character of the Coordination Functions Explanation	185
6.5 Conclusion	193
Chapter 7 - Applications	194
7.1 Introduction	194
7.2 Concepts and Conceptions	195
7.3 Identity Judgements.....	198
7.4 Special Ways of Thinking	206
7.5 Conclusion	216
Conclusion	217
Bibliography	218

Introduction

These are the things we may be said to have learned and to know. Yet, if I cease to recall them even for short intervals of time, they are again so submerged—and slide back, as it were, into the further reaches of memory—that they must be drawn out again as if new from the same place (for there is nowhere else for them to have gone) and must be collected so that they can become known. In other words, they must be gathered up from their dispersion. This is where we get the word *cogitate*... [T]he mind has properly laid claim to this word so that not everything that is gathered together anywhere, but only what is collected and gathered together in the mind, is properly said to be ‘cogitated.’

St. Augustine, *Confessions* X.xi.18

The topic of this essay is a problem concerning thought. The problem is presented by the fact that someone, having encountered or heard about something, and so having come to know something about it, may be rational in being disposed to use what she knows to infer something further, either about that thing in particular or about things in general. I call this *referential coordination*. Referential coordination is a way for identity to figure in thought. As such, it is closely related to what are known as *Frege cases*. But whereas Frege cases involve thoughts that the thinker rationally treats as being about different things, though they are in fact about the same thing, referential coordination involves a thinker having reason to treat her thoughts as being about the same thing. The problem of referential coordination is to explain what enables this that is absent in the other sort of case.

The kind of ignorance of identity exemplified in Frege cases has been subject to a great deal of discussion, and the philosophical literature dedicated to solutions and criticisms of those solutions is extensive. Referential coordination, by contrast, has received somewhat less attention. When it has, it is generally under the guise of addressing what is known as *coreference de jure*, which has more to do with some combination of the epistemic properties of deductive inferences and the semantics of coreferring expressions, such as anaphoric pronouns. This essay is partly based on making the case for a reversal of this situation: referential coordination should be treated as an independent problem, and there are reasons for thinking that the solution to it is explanatorily prior, in the sense that it constrains how those other problems should be solved. The case for this can

be made in the first instance by considering the nature of the problems. It also gains force from how it provides a relatively tractable explanatory test for competing theories of thought.

An exhaustive examination of such theories in relation to the problem of referential coordination is out of the question, so instead I will discuss three broad theoretical approaches prevalent in the contemporary philosophical literature: the appeal to Frege's notion of sense, as part of a theory of the contents of thought; the appeal to mental representations and ways of typing them, as part of a theory of the vehicles of thought; and the appeal to mental files, as part of a theory of how a thinker's information is organised. I will argue that representatives of each of these views face problems concerning the quality of the explanation of referential coordination that they can supply. In particular, I will show that they face problems with giving a non-circular explanation, deficits in psychological realism, and failure to cover a broad enough range of cases.

In light of these difficulties, I propose an alternative explanation that appeals to the normative functional properties of entities making up a minimal information-processing structure. This is an amended version of mental file theory that uses a distinction between mental files and file predications. File predications are the associations of mental files with bits of information. This is combined with an apparatus of defaults and defeaters familiar from entitlement epistemology. On this explanation, file predications have *coordination functions*, as I call them, that provide a default indication of sameness of reference, which is open to defeat by conflicting information. The rationality of the inferential dispositions can be understood as their coherence with the rest of a thinker's information, which is made available by the thinker's sensitivity to the reliability of information employed in the course of carrying out her projects and pursuing her aims.

The coordination functions explanation makes no use of a commitment common across all of the three views mentioned, also widely held in the background of much philosophical thinking about the mind. This is the idea that we should explain properties of thought by making use of *concepts*, understood in one way or another. At the same time, focusing on referential coordination can help make sense of what may be common ground between the many divergent theories of concepts. The coordination functions explanation is consistent with only some of the commitments made by these theories, and so also provides a way of evaluating the appeal to concepts to do theoretical work. It

also presents a way to reject *anti-explanationism*, the idea, associated with followers of Wittgenstein, that no substantive explanation of rationality is possible.

Referential coordination can be illustrated with the following simple case:

Lucy is introduced to Cicero at a party and is told that he is from Rome; after talking to Cicero for a while, she finds out various things, including that he is an orator; when asked later whether there are any Roman orators, she is able to answer ‘yes’.

In this case, Lucy thinks that Cicero is Roman and that Cicero is an orator, and on that basis she infers that *there is something* such that it is *both* Roman *and* an orator. In making this inference, the thinker treats those thoughts as being about, or referring to, the same thing. Treating one’s thoughts as referring to the same thing is manifested primarily in being prepared to conjoin the properties and existentially generalise so as to come to think the thought that something exists that has both. Actually making this inference—thinking the existentially general thought—is not so important, rather it is the disposition that is the mark of the thoughts being referentially coordinated. What makes it rational to be prepared to make this inference does not depend on the capacity to think general thoughts about what there is; when the disposition is present, it will engage this capacity, but what makes thoughts referentially coordinated is independent of it.

That the inferential disposition obtains does not present any particularly pressing problem by itself. There is a question of what might be the cognitive basis for the preparedness to make the inference, but this is a question that permits of answers at many different levels, almost all (though, crucially, not all) of which are well beyond the scope of this essay. The problem that the inferential disposition presents comes from the fact that the thoughts are treated as being about the same thing *rationally*. Talk of thoughts and inferences as rational can be understood in a few different ways. I mean something quite specific: the thinker has a reason to make that sort of inference. The inferential disposition is rational in that it is not a matter of states, episodes, or changes that a thinker finds unintelligible, that merely *happen* to her. Rather, it *makes sense* to her that she is prepared to make the inference. The problem is to account for this. What provides the reason she has for treating her thoughts as being about the same thing?

When the disposition is present and rational, thoughts are *coordinated* because there is a link between them, and they are coordinated *referentially* because that link concerns their referential properties. Hence I call the feature of thought in question *referential coordination*.

What is distinctive about referentially coordinated thoughts can be brought out by comparing the first case to a second, involving a slight change in the set up:

Lucy is told that Cicero is at the party and is told that he is Roman; after talking to Cicero for a while she finds out that he is an orator; when asked whether there are any Roman orators, she does not answer ‘yes’, because she did not realise that the person she was told about is the person she was talking to.

This sort of set up may be familiar from discussions of Frege cases. The similarity becomes more obvious if we suppose that Cicero introduced himself under an alternative name, say, ‘Tully’. Since she did not realise that Cicero is the same person as Tully (or, if you prefer, that the person called ‘Cicero’ is the person called ‘Tully’), then, were Lucy to be prepared to make the same inference she was in the first case, she would not be rational in doing so. Being a rational thinker, the disposition does not obtain; but her ignorance of the identity would prevent it from being rational were it to obtain.

We can compare the two cases in terms of the *referential content* of her thoughts, by which I mean the fact that she is thinking about the things she is thinking about specified just in terms of those things, in a ‘transparent’ way. The referential content of both thoughts is the same in both cases, but one and the same inference would be rational in the first case, and not in the second. We might describe the second case, quite naturally, by saying that she is using different *ways of thinking* about the same thing. In contrast, in the first case, she is using the *same* way of thinking. But what is it to use the same way of thinking? There must be something that provides a reason for treating the thoughts as being about the same thing which is present in the first case and which is absent in the second case. The fact that thoughts can be referentially coordinated or uncoordinated means that simply enumerating referential content is not sufficient for fully comprehending them; we need an account of ways of thinking for that.

This problem is a site of philosophical interest because referential coordination is needed for rational inferences of the sort that go towards forming the kind of rational conception of things of which creatures such as you and I are capable. Without it, this conception would be unrecognisable, perhaps impossible. Having a view of the world as populated by individuals, about which we can know that they have various combinations of properties and stand in various relations, and using this knowledge to formulate and see through out plans and projects, all in a form that hangs together for us in a coherent and plausible way—this is not something we could do without, epistemically limited creatures as we are. Whatever its wider implications, understanding the part of our lives that rational thought occupies requires understanding ways of thinking, and so requires an explanation of referential coordination.

Chapter one, ‘The Problem of Referential Coordination’, develops the problem in more detail. Referential coordination involves an *indication of sameness of reference* that cannot consist in an explicitly formulated identity, but must instead be *inferentially basic*. The problem is that it not obvious what this might be. Anti-explanationists will be sceptical that any explanation can be given, since it requires a *personal-level* explanation, and nothing at the personal-level seems available; this is correct about what is required, but the claim about what is available is too quick (§1.2). I then provide reasons for treating referential coordination as independent of some related problems discussed in the literature, including Frege cases. I also present reasons for treating referential coordination as a prior problem, in the sense that the best explanation of it should constrain solutions to those other problems rather than the other way around (§1.3). I then set out some general theoretical questions about thought, and show how they supply dimensions across which different responses to the problem of referential coordination may differ. A sketch of the coordination functions explanation is provided to show how it fits into this framework (§1.4). I end by giving a formulation of the explanatory criteria to be employed in the argument in later chapters (§1.5).

Chapter two, ‘Concepts’, discusses the relationship between referential coordination and debates in the theory of concepts. A problem in this area is that the question ‘what are concepts?’ is often given distinct answers that are not always clearly competitors (§2.2). Introducing referential coordination helps with this: the divergent answers to the question can either be seen as differing over the question of what explains referential coordination, or else as providing a way of describing patterns of referentially coordinated thoughts (§2.3). This answer does better than some alternatives at providing a means of

achieving a degree of dialectical unity (§2.4). I then apply this to a claim made by Machery, to the effect that philosophers and psychologists talk about different things when they talk about concepts. The distinction is a useful one, though how to evaluate his claim is contingent on having a view on what the best explanation of referential coordination is (§2.5). More generally, the correctness of a commitment to the existence of concepts, at least in the explanatory sense, is itself contingent on a particular sort of explanation of referential coordination being correct. If the explanation to be defended here is the best, then this sort of explanation is incorrect, and so the theory of concepts rests on shaky foundations (§2.6).

Chapter three, ‘Sense Theory’, evaluates explanations based on a development of Frege’s notion of sense as a theory of the content of thoughts. I begin by setting out the central tenets of sense theory in relation to the sense-theoretic explanation of referential coordination, and discuss some disagreements between those who accept Frege’s notion of sense relevant to the explanatory problem (§3.2). I then present an initial challenge to the sense-theoretic explanation: it needs to supply a theory of what it is for particular senses to feature in the contents of thoughts (§3.3). I examine some paradigm treatments of sense theory, presented by Chalmers and Peacocke, to see what responses to this challenge they can provide (§3.4). I argue that the response to the challenge that these treatments can provide is at best circular, and show how this problem can be traced back to one of the central commitments of sense theory in such a way that it can be generalised beyond the particular treatments examined (§3.5).

Chapter four, ‘Typed Representations Theory’, evaluates explanations based on the claim that thought involves mental representations that stand in relations of type-identity. Typed representation theory is best treated as an alternative to sense theory (§4.2), though it faces a similar challenge: it needs to substantiate the relevant notion of a type (§4.3). The theory finds a clear expression in the language of thought hypothesis, closely associated with Fodor. Proponents of typed representations theory often appeal to the idea that mental representations are expressions in the language of thought, with the relevant types being types of expression in that language. However, the syntactic typology Fodor endorses is lacking in explanatory content sufficient for meeting the challenge (§4.4). More substantive proposals, owing to Sainsbury and Tye (§4.5) and Prinz (§4.6) can be shown to suffer from various explanatory deficits. Millikan’s related discussion of what she terms ‘grasping sameness’ fails to respect the fact that referential coordination is

a rational and not purely dispositional phenomenon, a mistake that the appeal to mental representations is apt to share (§4.7).

Chapter five, ‘Mental File Theory’, evaluates explanations based on the idea that there are mental entities that serve to organise a thinker’s information. I offer a reconstruction of how this might be motivated, on which mental files are stipulated to be the basis for the paradigmatic inferential dispositions, which can then be made the subject of substantive claims that explain the rationality of those dispositions (§5.2). What I call *basic model explanations* bring in some further capacity of the thinker, directed at her mental files, to provide the explanation. I discuss two basic model explanations that can be gleaned from the literature, due to Schroeter and Lawlor, and show that they require thinkers to have cognitive capacities that are psychologically unrealistic (§5.3). More robust explanations use more robust claims about what mental files are. Recanati has defended such a model, on which mental files exhibit properties similar to indexical expressions in public language, such as ‘I’, ‘here’, and ‘now’. Some discussion is needed to see what plausible explanation this model might provide (§5.4). The most plausible explanation fails due to the failure to be sufficiently general; there are cases that are not plausibly treated as involving the kind of mental file that makes the explanation applicable (§5.5).

Chapter six, ‘Coordination Functions Theory’, presents and defends the theory of referential coordination advocated in this essay. I elaborate the coordination functions explanation at length, and raise and respond to some objections. Particularly important is the point that, though coordination functions do not need to be based on mental files, it is a part of its being the best explanation that the appeal to coordination functions is given a mental file-theoretic basis (§6.2). The coordination functions explanation avoids the explanatory deficits associated with the competing explanations, a conclusion which I bolster by providing some comparisons with those theories. This provides the main argument in favour of the coordination functions explanation (§6.3). I then draw out the consequences of accepting it by elaborating on the character of the explanation. I show how it answers the three questions introduced in the first chapter, and elaborate on how it offers a particular take on the rationality of the inferential dispositions as a matter of coherence between inferential dispositions and the thinker’s broader stock of information, rather than one of information that excludes possibilities. This provides points of overlap with the basic model explanations looked at in chapter five, and brings out the relatively deflationary character of the coordination functions explanation (§6.4).

Finally, chapter seven, 'Applications', completes the argument for the coordination functions explanation. I close the discussion of concepts by spelling out how coordination functions provides a reason to endorse the eliminative stance introduced in chapter two. This conclusion is aided by adapting a distinction between concepts and conceptions, where conceptions are to be thought of as bundles of information linked by coordination functions. Conceptions can account for much of what concepts are traditionally held to do (§7.2). The notion of a conception can be used to apply the coordination functions explanation to identity judgements (§7.3) and ways of thinking at issue in so-called demonstrative and indexical thoughts, albeit one that involves a reduction in the theoretical ambition. Though some with more inflationary ambitions might see this as a reason to object, I argue that conceptions provide a way of accommodating those ambitions in line with the deflationary character of coordination functions theory (§7.4).

Chapter 1 - The Problem of Referential Coordination

1.1 *Introduction*

This chapter develops the problem of referential coordination in more detail, and provides an overview of what might be involved in solving it. The aim is not to advocate any particular position on referential coordination, but to get clear on what is at stake between different explanations, and why it is worth seeing which is best.

Referential coordination presents an explanatory problem because it is a way for identity to figure in thought that cannot be assimilated to explicit belief in one thing being identical with another. With referentially coordinated thoughts, identity figures in an *inferentially basic* way, and this prompts the need for some theoretical ingenuity in specifying what this could involve. I discuss the *anti-explanationist* position that taking referential coordination as an explanatory problem is somehow mistaken. This might be motivated on the basis of restricting the explanation of rational phenomena such as referential coordination to the *personal-level*. While this motivation is faulty, it does help to articulate an important aspect of what an explanation of referential coordination needs to do (§1.2).

The philosophical literature features extensive discussion of a number of related problems that bear on referential coordination, including those raised by Frege's seminal discussion of identity and proper names. I argue that there are good reasons to treat referential coordination as an independent and prior explanatory problem, in that the best explanation of referential coordination provides a constraint on how to deal with these other problems (§1.3). I then set out some general questions and three theoretical approaches represented in the literature that I will discuss in subsequent chapters, and provide a sketch of the *coordination functions* explanation that I will defend, and explain how it differs from the competing explanations in terms of the general questions. (§1.4). I end by setting out the explanatory criteria that will form the substance of the argument in later chapters (§1.5).

1.2 *Referential Coordination as an Explanatory Problem*

This section expands on the problem of referential coordination, and introduces some terms that will be useful in discussing proposed solutions.

1.2.1 What is referential coordination?

In the introduction, I used the following example to illustrate the problem of referential coordination:

Lucy is introduced to Cicero at a party and is told that he is Roman; after talking to Cicero for a while she finds out various things, including that he is an orator; when asked whether there are any Roman orators, she is able to answer ‘yes’.

It is not hard to come up with cases that exemplify this same pattern: an individual has some thoughts (two will do), and puts them together inferentially to rationally support a third. In such cases, a thinker has referentially coordinated thoughts. The problem is to explain what makes this the case.

The thoughts in these cases, it should be made clear, are thoughts in the sense of standing psychological states, of the kind often given the label of *propositional attitudes*. Thoughts in this sense are states that we can report by using an assertoric sentence in the context of a psychological verb, paradigmatically *believing*, though arguably, *knowing* is more central (cf. Williamson 2000). Since knowledge has the implication of the truth of what is known, it may be convenient to go with the more hedged notion of belief. I will just call them thoughts. These are not what might be considered *episodes* of thinking, or *occurrent* thoughts. Rather than imagined acts of speech in inner monologue, these thoughts are what such imagined acts would express. Nor are they episodes or acts of deciding or judging that something is the case, consciously or otherwise. They are employed in inference, but they are not themselves things that occur or happen. Similarly, inferences can be thought of primarily as relations of support between thoughts, not acts or episodes of inferring.

We can describe the original case by making the psychology of the thinker more explicit like this:

- (1) Lucy thinks that Cicero is a Roman;
- (2) Lucy thinks that Cicero is an orator.

(3) On the basis of (1) and (2), she infers that there is something that is a Roman and an orator.

When thoughts like those reported in (1) and (2) obtain together, the thinker will have a particular sort of inferential disposition. This is the disposition that is manifest in the inference as described by (3). The disposition is a preparedness to *conjoin and generalise*. The thinker infers that there is something (existential generalisation) that has both properties (conjunction).

Conjoining and generalising in this way is, in truth, not a particularly interesting inference. There are more interesting kinds of inference that demonstrate the importance of referential coordination just as much. For the most part, these are *ampliative* inferences, those in which the thinker brings out more than the merely deductive consequences of what she thinks. These involve more than just the introduction of conjunction and existential generalisation. They include the subsumption of individuals under general categories (Lucy believes that Roman orators are usually loquacious bores, and so believes that Cicero is a loquacious bore), inferences that guide action (Lucy believes that loquacious bores are not worth talking to, and so makes her excuses and talks to someone else), and the revision of previously held beliefs (Lucy finds out that Cicero is an engaging person to have a conversation with and so changes her mind about Roman orators). However, the less interesting disposition to conjoin and generalise will be present whenever the more interesting dispositions are present, and it brings out particularly clearly what is at issue. I will take the inferential disposition to conjoin and generalise as the paradigm case, and refer to it as the *paradigmatic inferential disposition*.

The rationality of inferential dispositions of both the paradigmatic kind and the other more interesting kinds depend on having a *reason* to treat two thoughts as about the same thing - or else, having a reason not to do this. The nature of this reason has to go beyond simply the fact that, as it happens, they are about the same thing. To use a visual metaphor that comes easily to hand, we cannot *see* what our thoughts are about independently of our thinking them. Rationality under such a limitation is about having reasons for taking our thoughts to represent the world in a particular way, as reflecting identity relations that really do obtain between what things there are. To stretch the visual metaphor further, we might say that a thinker has a *perspective* on the referential content of her thoughts that can be articulated in terms of the paradigmatic inferential

dispositions and the reasons that bear on them. This perspective is a crucial part of being a rational thinker.

This points to how, at its heart, referential coordination bears on how information can be combined so as to extend a thinker's knowledge (or, as it might be, justified belief). It is a plausible idea that the paradigmatic inference in the absence of referential coordination could not be knowledge-extending even though truth-preserving, since it would be a process of bringing about beliefs irresponsibly, lacking the necessary circumspection that distinguishes knowledge from stabs in the dark. If that is disputed, then the point stands that it would be knowledge-extending only in an attenuated sense, as it would fall far short of an epistemic ideal. It would not be the activity of an epistemically responsible inquirer. Ampliative, action-guiding, and revisionary inferences again provide for more interesting kinds of case on this score. Even so, the point about inferences extending knowledge is just as pressing for the less interesting paradigmatic inferences that provide the clearest expression of the phenomena.

Referential coordination means that something is needed in addition to referential content in order to understand how thought works. To employ the phrase that gives the title to this essay, thought is not simply *about* things, but involves *ways of thinking* about things. Another phrase often used to express the same idea is that *thought is more fine-grained than reference*, the idea being that there can be a many-to-one relationship between what we think about and how we think about it. In the case of referential coordination, the rationality of the paradigmatic disposition to conjoin and generalise involves the thinker employing the *same* way of thinking about the thing in question.

A slight modification to the example used above gives us an example where this is not so:

Lucy is told that Cicero is at the party and is told that he is Roman; after talking to Cicero for a while she finds out that he is an orator; when asked whether there are any Roman orators, she does not answer 'yes', because she did not realise that the person she was told about is the person she was talking to.

In this example, the same inferential disposition would *not* be rational even though the referential content of the thoughts is the same. The thinker is employing *different* ways of thinking. As I said in the introduction, this kind of set up may be familiar from

discussions of what are known as Frege cases, which I will discuss in §1.3 below. Before then, I want to say more about why referential coordination presents an explanatory problem.

1.2.2 Why does referential coordination present an explanatory problem?

Where thoughts are referentially coordinated, there must be something that provides an *indication of sameness of reference* that is sufficient for making the paradigmatic inferential disposition rational. This sort of indication would be absent when thoughts are not referentially coordinated. Talk of sameness in ways of thinking is really only a label for the presence or absence of indications of this kind. The problem is that it is not at all obvious what the indication of sameness of reference might be.

Whatever it is, it cannot be an additional thought that figures as a further premise in the inference. Such a premise would, presumably, need to concern the identity of the thing referred to by each thought; call this an *identity premise*. In the case of Cicero at the party, the identity premise might perhaps be expressed by Lucy by saying ‘Cicero and the guy I was talking to are one and the same person.’ This cannot be a solution to the problem. For one thing, bringing in an identity premise would represent something quite different from simply thinking of something in the same way; it would stand for the realisation that what was thought of as two things is in fact one thing. But the paradigmatic inference does not involve this sort of realisation.

That point aside, bringing in an identity premise does not work because it merely multiplies the problem. The identity premise would itself need to be referentially coordinated with the other thoughts. If Lucy thinks Cicero is a Roman, and the guy she was talking to is an orator, then the identity premise expressed by ‘Cicero and the guy I was talking to are the same’ must be referentially coordinated with *those* thoughts in order for that inference to be rational. The thing to be explained would then simply occur again, twice over. As Campbell puts it, the sort of inference that is the mark of referential coordination *trades on sameness of reference* (1987/88: 278-9; see also Campbell 1994; cf. Sainsbury 2002: 133-6, Recanati 2012: 47-50).

The indication of sameness of reference sufficient for explaining referential coordination, whatever it is, must therefore operate in a wholly different manner to explicit belief in one thing being identical to another. We can put this by saying that the indication is

inferentially basic, in the sense that it is not just a further premise in the inference but more like a property of the thoughts that constitute its premises. This makes referential coordination particularly puzzling, and why the problem of referential coordination is a real *problem* and interesting in its own right, quite apart from its crucial place in our cognitive lives. Explaining it seems to require some theoretical ingenuity, since there is no obvious candidate for what an inferentially basic indication of sameness of reference is

One might think that the appearance of a problem here is simply a product of how I have chosen to represent the cases where the inferences are rational and those where they are not. The *form* of the inferences in either case—the one Lucy is disposed to make, and the one she is not disposed to make—can be represented in a more perspicuous way using the standard notation of first-order predicate logic. In the first case, the form of the inference can be perspicuously written out like so:

$$\begin{array}{l} (1) F(a) \\ (2) G(a) \\ \hline (3) \exists x (F(x) \& G(x)) \end{array}$$

while in the second case, the form of the corresponding inference can be more perspicuously written out like so:

$$\begin{array}{l} (1) F(a) \\ (2) G(b) \\ \hline (3) \exists x (F(x) \& G(x)) \end{array}$$

These more perspicuous representation present an obvious difference between the two cases: a schematic letter of the same type, ‘a’, appears in the two premises of the first inference, whereas it appears only in one premise of the second. The form of the two inferences is therefore different. One is a *valid* inference, and the other not. The first inference is an instance of a truth-preserving form, whereas the second one is an instance of a form that does not preserve truth. Lucy, being a rational thinker, is disposed to engage in valid inferences and not in invalid ones. Once this difference in validity is noted, the thought goes, the problem goes away.

But representing the form of the inferences like this does not make the problem go away; it just provides another way of illustrating the problem. The more perspicuous representation of the inferences with the same letter for the object in the first and different letters in the second *represents* the fact that there is a difference between the inferences that can be represented, and this is just what calls for explanation. What makes it the case that the inferences in the two cases are better represented in one way or the other? Merely adverting to the way in which the inferences can be represented leaves this question untouched. The type of letter used to represent the inferences cannot literally be what provides the indication of sameness of reference. If some analogue of schematic letters in the notation of first-order predicate logic is what provides the explanation, then this could only be a substantive hypothesis in significant need of elaboration and support.

The problem of referential coordination does not simply concern, and is not solved by accounting for, the validity or otherwise of inferences the thinker is disposed to make. What is being sought is an explanation of the rationality of the thinker's disposition to engage in the paradigmatic inference. That an inference is valid may be part of the reason a thinker has to make it, and in other cases, invalidity may be part of the reason she has not to make it. But it is not the validity or lack of validity itself that matters. It is the presence of an indication of sameness of reference, where this is an indication that the inference would be valid. The problem is to say what form this indication takes. The rationality of inferential dispositions, in the relevant sense, and the validity of inferences therefore must be held apart, despite their obviously close relationship.

1.2.3 Anti-explanationism and the personal-level

I have claimed that referential coordination involves an inferentially basic way for identity to figure in thought, and that this means that it presents an explanatory problem. It is possible to treat this claim with scepticism. According to what I will call *anti-explanationism*, there only appears to be a problem, and that appearance perhaps points to a mistaken philosophical approach to thought. Ultimately, the best way to refute anti-explanationism is to provide an adequate explanation. But it is useful to consider the possible motivations for it, as it helps to articulate a constraint on what an adequate explanation needs to do.

Anti-explanationism about referential coordination would require rejecting the general principle that what seems to call for explanation can indeed be explained. Instead, the thought would go, we would need to move away from whatever it was that made it seem that an explanation was required. This attitude is sometimes given the label *philosophical quietism*; it is opposed to the idea that philosophical problems call for constructive theories. Proponents of quietism, such as McDowell (1994), sometimes find the origins of this attitude in the anti-theoretical animadversions of the later Wittgenstein:

[T]he clarity that we are aiming at is indeed complete clarity. But this simply means that the philosophical problems should completely disappear.

The real discovery is the one that enables me to break off philosophising when I want to. – The one that gives philosophy peace, so that it is no longer tormented by questions which bring itself in question...

There is not a single philosophical method, though there are indeed methods, different therapies, as it were. (1953: §133)

This kind of position is often based on identifying the 'felt need' for explanation as an over-enthusiasm for the theoretical style of the natural sciences, a hangover from the scientific revolution, or else driven by some need to relieve some deeper philosophical 'anxiety'. Quietism involves a change in deeply held views about philosophical questions, particular those concerning the status of ourselves as rational beings, and what it means to come to understand ourselves as such. Indeed, philosophical quietism might be expressed as the thought that philosophical questions, or rather the felt need to pose them and provide substantive theories by way of response, are themselves based on a mistaken attitude to what those questions are about.

Evaluating the independent attractions of philosophical quietism would go far beyond the scope of this study, requiring, at a minimum, a discussion of Aristotelian metaphysics, the development of the idea of disenchanted nature and disengaged reason, the history of the Scientific Revolution, Hegelian Absolute Idealism, Deweyan pragmatism, and other such mind-boggling topics (see Pippin 2002 for related discussion). These aside, the general thesis of philosophical quietism can be brought back to earth by noting how it is open to refutation, or at least qualification, by specific

counter-examples (cf. Dummett's 1978 remarks on Wittgenstein). Whatever appeal it may have, there needs to be some specific reason to think that it gets witnessed by particular phenomena. What might that be in the case of referential coordination?

Anti-explanationism could be independently motivated by arguing along the following lines. The cases that apparently raise the explanatory problem are those in which a person undergoes certain transitions in their thoughts. They are undergone by a thinker, for reasons that the thinker has. They therefore call for *personal-level* explanation. This means that it would not be *genuinely* explanatory merely to refer to something *sub*-personal, such as what is going on in a part of their computational psychology, or in their neurology, or whatever. Exactly what counts as being a personal-level explanation is not wholly clear. A helpful suggestion (Hornsby 2000) is that it is one that involves properties of a person, rather than properties or states of things that are not persons, though may be parts of a person, or otherwise intimately related to a person. This is admittedly still somewhat in need of clarification, but we may accept it as a guide.

Dennett, when introducing the (or rather, an initial version of the) distinction between personal and sub-personal levels of explanation in the context of the phenomenon of pain, writes:

Abandoning the personal level of explanation is just that: *abandoning* the pains and not bringing them along to identify with some physical event. The only sort of explanation in which 'pain' belongs is non-mechanistic; hence no identification of pains or painful sensations with brain processes makes sense... (Dennett 1969: 94)

In Dennett's hands, the point of the distinction is to reveal how some explanatory projects involve a problematic change of subject. The distinction provides the resources to pinpoint what is going on in cases where an explanatory proposal leads to darkness rather than illumination; when the phenomenon in question concerns persons and their reasons, changing the subject to talk of impersonal mechanisms and their dynamics would be a philosophical error rather than the beginning of wisdom (though see Dennett's remarks about Wittgenstein and Ryle at op. cit.: 95-6).

The thought might then be that, if the explanation needs to involve personal-level facts, the only relevant facts concern what the person, the thinker herself, could express and

recognise as what it is that makes sense of her inferential dispositions. But if so, the only thing they could say, if they found it sensible to say anything at all, would be “what I was thinking about was one and the same thing.” At that point, we hit bedrock (Wittgenstein 1953: §217) where referential coordination is concerned.

This forms the basis for two worries that point to anti-explanationism. It might be thought that hitting bedrock presents a perfectly adequate explanation of the inference that person made in terms that make sense to them, once we know that they would say “what I was thinking about was one and the same thing.” We as theorists can make sense of their thinking that it is the same thing in just the same way: were *we* in the same position, we would say exactly what they said, and it would make sense for us to do so just as well. All the explanation we need is there in what the thinker says. To think that anything further is needed in the way of explanation is to confuse the phenomenon of a person thinking in a rational way with something else. So the idea that providing an explanation is in any real sense a problem simply lapses in the face of a proper appreciation of what it is that needs explaining and how. Call this the *levels* worry.

The same thought might be pushed to a more radical conclusion. If referential coordination requires personal-level explanation, the explanation must cite reasons that depend on properties of the person, the thinker herself, and not just on some part or ‘internal’ mental mechanism. The normal way for such reasons to play a role in a person’s life is to be some proposition that a thinker can express. But if the *only* reason that the thinker could express is “what I was thinking about was one and the same thing,” then that reason provide no explanation at all. In which case, it might seem that the demand for an explanation cannot be discharged. So if we accept that there is a problem, then we must also accept that there can be no solution. That is a paradox that should lead to a rethink of how the phenomenon is being conceived; call this the *radical levels* worry.

These worries support anti-explanationism in different ways. With the levels worry, the thought is that it is *obvious* what is going on. With the radical levels worry, the thought is that it is only not obvious what is going on, and so there only *appears* to be a genuine problem, if one thinks of the phenomena in such a way as to make an explanation impossible. To vary the example a bit, suppose Lucy looks out of the window and sees a red car, and later notices that it is a Mercedes; Lucy is rationally disposed to infer that there is something that is red, and is a Mercedes. According to the levels worry, it is not mysterious what is going on because if asked, Lucy would say that she was thinking

about the same thing; we can imagine ourselves being in the same situation, and we would be disposed to make the same response. According to the radical levels worry, the response is no explanation, but there only appears to be an explanatory problem if we think that something more is needed.

The two worries diverge over whether the thinker's avowal that "what I was thinking about was one and the same thing" really is an explanation. On this point, we should side with the radical levels worry. The fact that a thinker might be willing to make that statement cannot be the reason that a thinker has for treating her thoughts as about the same thing. The statement merely asserts the identity of what is thought about with itself, and that cannot be the reason, because the identity of the thought about individual is the same in both the case of referential coordination and its absence. Moreover, the statement is only an *expression*, and not an explanation, of referential coordination. We knew already that they had referentially coordinated thoughts, and of course this can be expressed by saying "what I was thinking about was one and the same thing." But *why* do they say that? The levels worry gives no reason to think that this question is ill-formed. It is really implicitly offering an explanation rather than showing that no explanation needs to be given.

The problem with the radical levels worry is connected to this. The assumption made is that personal-level explanations of rational phenomena such as referential coordination could *only* be given in terms of reasons a thinker can express. These are what we might call her *discursive* reasons (cf. McDowell 1994: 163-166). The worry put in these terms is that there are no appropriate discursive reasons and so there can be no solution to the alleged problem. The thinker can only give expression to the phenomenon, not cite anything more basic that could figure in an explanation. And so the conclusion could be drawn that the expression of referential coordination is all we could ask for. In which case, it might seem better to reject the idea that there is a real problem at all. But while it is reasonable to think that there are no appropriate discursive reasons, that is no grounds for pessimism about explaining referential coordination.

One might grant that the paradigm of personal-level explanation is that of a linguistically expressible reason. It does not follow that only reasons of that sort may figure in such explanations. Facts about the thinker independent of her capacities for linguistic expression can be explanatory of the rationality of her thought. One may have a sensitivity to the particulars of a situation and what it calls for that goes beyond what one

can express at the time or afterwards. For example, a skilled bridge player may be able to bid and play her cards in just the right way to make the contract, although she is unable to say in advance exactly what she will do, or spell out afterwards what it was that called for bidding and playing the cards in the way she did. It was not, for all that, a merely automatic response to the situation, but something she did that made sense to her; she played the cards in the way she did because she was trying to make the contract and because that was the way to do it. It is not some extraordinary mystical awareness that allows her to respond rationally, but a capacity of a perfectly ordinary kind that one may have, perhaps as a result of training or native talent. The problem would be to explain that capacity. The restriction to discursive reasons anyway jibes badly with the basic nature of referential coordination, and it is unlikely that the explanation will bottom out with our linguistic capacities.

While the levels worry can be rejected as implicitly offering a faulty explanation, the radical levels worry calls for a more nuanced response. *If* there is an explanation to be had, it must be an adequately personal-level explanation that involves something other than discursive reasons. For the time being, we have no reason to think that there is not an explanation to be had, so the anti-explanationist idea that no explanation of referential coordination should be sought cannot be made particularly pressing on the basis of unavailable discursive reasons. Rather than grounds for scepticism, the point about the absence of discursive reasons shows rather another way in which referential coordination presents a real problem.

1.2.4 Summary

Referential coordination is a feature of thoughts that sustains our rational conception of things. An explanation of referential coordination needs to say what it is that provides the indication of sameness of reference in cases where it is rational to perform a paradigmatic sort of inference, from thoughts ascribing properties to things to an existential generalisation conjoining those properties. This indication cannot consist in more obvious ways for identity to figure in thought (explicit or linguistically expressible beliefs in identities), and so presents an explanatory problem that calls for some theoretical ingenuity.

1.3 Related Problems

There are a number of other phenomena closely related to referential coordination that have been more widely discussed. These include what are known as Frege's puzzle of identity, Frege cases, and coreference *de jure*. There is also the more general issue of how to understand intentionality, the fact that thoughts are about things, or have referential content, at all. This section provides enough detail on these problems to motivate the idea that referential coordination should be treated as an independent problem that is explanatorily prior, in that the best explanation of referential coordination provides a constraint on how to deal with these other problems. At the same time, in some cases the relationship is close enough that they exert some pressure on what counts as a good explanation of referential coordination.

1.3.1 Frege's puzzle of identity

Frege's seminal essay 'On Sense and Reference' (1892/1948) raises a number of problems that indicate the insufficiency of referential content for comprehending thought. There are at least two problems, one concerning thoughts about identity, the other concerning rational though inconsistent propositional attitudes, that are useful to hold apart. I will discuss the puzzle of identity first.

Frege's discussion begins by raising a puzzle concerning identity. The puzzle comes down to how to understand properties of sentences that feature noun phrases surrounding the identity sign. A sentence of the form 'a=a' seems to be *uninformative*, or at least knowable *a priori*, since all it says is that something is identical with itself; everything is identical with itself; so no possibilities are ruled in or out (except, perhaps, that a exists). However, a sentence of the form 'a=b' *can* be informative, and can express some piece of empirical information. But if an 'a=b' sentence is true, then it is true if and only if the 'a=a' sentence is true. In which case, they should not differ in informativeness if their meaning is the same; they do differ in informativeness; so there must be some difference in the meaning of the two sentences, on some understanding of 'meaning'.

Frege used this puzzle about identity to motivate a distinction between the reference of a referring expression and its sense:

It is natural, now, to think of there being connected with a sign (name, combination of words, letter), besides that to which the sign refers, which may be called the referent of the sign, also what I would like to call the sense of the sign, wherein the mode of presentation is contained.... The referent of "evening star" would be the same as that of "morning star," but not the sense. (op. cit.: 210)

Difference in sense is what is supposed to explain the informativeness of, for example 'Cicero is Tully', as compared to sameness of sense with the uninformative 'Cicero is Cicero'.

Although Frege puts the puzzle about identity in terms of sentences, it can also be considered as a problem about a particular kind of thought (this distinction can be somewhat blurred by those who extend Frege's semantic programme to the idea that the *contents* of thought are senses of sentences). Since the present topic is one that concerns thought, we can set aside the puzzle as concerning sentences and instead concentrate on what are known as *identity judgements*. Identity judgements are explicit beliefs in the identity of one thing with another, which might be expressed in sentences of the form 'a=b' or 'a is the same thing as b'. Frege's discussion raises two questions about identity judgements:

How can identity judgements have empirical significance? Everything is self-identical, and so we know automatically that one thing is identical to itself, and not identical to everything else. But then how is discovering an identity possible, as it evidently is?

How is there a difference in the rational role of identity judgements? Trivial identity judgements are obvious, perhaps knowable *a priori*, whereas informative identity judgements can only be knowable on the basis of empirical discovery. What accounts for this difference?

These questions can be understood in terms of the loose notion of ways of thinking introduced above. The empirical significance and rational role of identity judgements is dependent on an epistemic achievement that changes one's ways of thinking; as I put it in §1.2, they involve *realising* that what one thought of as two things is in fact one and the same thing. So the empirical significance of an identity judgement can be thought of as

depending on having thoughts that exhibit different ways of thinking. Similarly, merely entertaining identities, as opposed to believing or ‘judging’ them, can be understood as entertaining a potential change to one’s distinct ways of thinking. Understanding how identities can be discovered and brought into a thinker’s stock of information, and as a result modify the inferences that she can rationally make, requires first understanding the inferences that she can rationally make before that discovery.

Answers to these questions might therefore be given in terms of the more basic way in which the identity of what we think about is represented in thought, which is referential coordination. Sameness in ways of thinking comes before operations on distinct ways of thinking. To get a workable theory of the latter, it would be a good idea to get a satisfying explanation of the former. And if that’s right, then we should treat referential coordination as a problem that is prior to that of answering the two questions about identity judgements. This does not mean that providing with a good explanation of referential coordination is *wholly* independent of dealing with identity judgements. A good explanation of referential coordination should ideally extend to an account of identity judgements. If these are to be understood as modifying patterns of referential coordination, then an explanation of referential coordination is better if it can be extended to also provide an account of the empirical significance and the rational role of identity judgements.

A short digression on terminology: as Frege puts it in the quotation above, the sense of an expression is that “wherein the mode of presentation is contained.” Following Frege’s discussion, the informativeness of identity statements and the rationality of believing incompatible propositions are often discussed in terms of modes of presentation. Different modes of presentation are contained in the sense of the expressions, or different modes of presentation are involved in the thinker’s attitudes to the incompatible propositions. The ease with which this terminology can be used can obscure the fact that it is not always clear exactly what a mode of presentation is supposed to be. Sometimes, it seems to be similar to how I have been using ‘way of thinking’, as a stand-in for *whatever* it is that does the explanatory work. Sometimes, they seem to figure as a particular explanatory proposal, tied to Frege’s notion of sense.

Understanding what is intended by the term ‘modes of presentation’ requires some sensitivity to whether it is being used to outline a problem, or to propound a solution to a problem, which is not always a straightforward matter. I want to suggest that, where

mention is made of modes of presentation, this may be reinterpreted so as to make the link to referential coordination more apparent. Sameness of mode of presentation can be understood as just whatever it is that provides an indication of sameness of reference, and difference in mode of presentation can be understood as the absence of such an indication. In the technical sense, as meaning something akin to Frege's notion of sense, sameness and difference in mode of presentation stands for a substantive proposal about referential coordination that is subject to explanatory disconfirmation. Interestingly, in his *Varieties of Reference*, Evans (1982: 18-22) makes explicit a preference for cashing out the notion of Fregean sense in terms of ways of thinking, though he intends it to serve as a term for "what makes it the case that a subject's thought is a thought about the object in question," (op. cit.: 20; see also Evans 1985: 294) thus giving it in a much more substantive meaning than the one I am adopting here.

1.3.2 Frege cases

Frege employed the distinction between sense and reference to provide treatments of a number of other problems in semantics, including reference to non-existents and reference in indirect discourse, including what are now known as Frege cases. Frege cases are those that fit the pattern of the example used above to provide the contrast case to cases of referential coordination: a thinker has thoughts about the same thing but does not treat them as about the same for purposes of rational inference. Frege's well-thumbed example of the morning star and the evening star provides us with the following case:

Lucy looks up and sees the brightest star in the morning sky and calls it 'Phosphorous'. She thinks that Phosphorous is a planet. Later, she looks up and sees the brightest star in the evening sky and calls it 'Hesperus'. She thinks that Hesperus is not a planet. Hesperus and Phosphorous are in fact both the planet Venus.

Such cases are sometimes discussed in the context of arguments about the meanings of propositional attitude reports. The problem there is to provide an analysis of, for example,

- (i) Lucy believes that Phosphorous is a planet.
- (ii) Lucy believes that Hesperus is not a planet.

that allows for both (i) and (ii) to be true even though Hesperus is identical to Phosphorous and Lucy is not simply being irrational.

Frege cases can also be looked at in terms of the thoughts themselves rather than sentences about them. Understood that way, a Frege case involves a thinker with thoughts whose referential content is the same but are not referentially coordinated. The fact that referential content is the same shows that we need something other than referential content to comprehend thought. The problem is essentially that of accounting for what makes it the case that there are two distinct bundles of referentially coordinated thoughts. This point can be obscured by the presentation of Frege cases as only involving two thoughts which the thinker does not treat as about the same thing; but where this is one thought, there will almost always be many more, and even when there is only one there *could* be more in the bundle. If that's the right way to think about the relationship between referential coordination and Frege cases, then it is a good idea to have an account of what makes for there being a bundle in order to get a sense of what might account for their being distinct.

Frege cases are always also cases of referential coordination, though not *vice versa*. So we should treat them as independent problems. Frege cases also involve the *absence* of referential coordination. It may well be that there are *multiple* explanations of what is going on in a Frege case, depending on what is the correct way to think about referential coordination. It is not a given that Frege cases present an explanatorily unified phenomenon. By contrast, it seems more likely that referential coordination presents an explanatorily unified phenomenon. Getting a grip on Frege cases requires first understanding what is absent in them. Because of this, referential coordination should be treated as explanatorily prior. Although thoughts for which the thinker is ignorant of the identity of referential content provide a comparison helpful for setting out what referential coordination is, focusing on them will not necessarily provide a means of getting at what is going on (compare Rattan 2013, who argues for a similar conclusion from within the Fregean perspective). For these reasons, referential coordination should also be treated as explanatorily prior to Frege cases. The explanation of the latter should be constrained by the best explanation of the former.

There are some cases that can be seen as variations on classic Frege cases that have played no small part in philosophical discussion about thought and language that are worth discussing in relation to referential coordination. One variation that has received a very

large amount of attention is related to some cases Frege discussed in a later essay, 'Thought' (1918/1956). Perry's (1979) discussion of indexical thoughts has provided a rich seam of cases involving thinkers holding incompatible attitudes to propositions involving themselves and their current temporal or spatial location. Perry's first case makes the point:

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 (op. cit.: 3)

As it is usually put, Perry believed that the shopper with the torn sack was making a mess, but he did not believe that *he himself* was making a mess, even though he was the shopper with the torn sack.

Similar to this is a case discussed by Evans, involving what are taken to be examples of demonstrative thoughts:

Suppose a person can see two views of what is in fact one very long ship, through two windows in the room in which he is sitting. He may be prepared to accept 'That ship was built in Japan' (pointing through one window), but not prepared to accept 'That ship was built in Japan' (pointing through the other window). (1982: 84)

The special detail in this case is it seems that different thoughts are expressed with exactly the same linguistic expression ('that ship') that work in such a way so as to have the same reference.

Another variation is demonstrated by some cases presented by Kripke (1979) as part of his discussion of beliefs made available via understanding, or misunderstanding, uses of proper names. One case involves beliefs expressible by inter-translatable proper names in different languages:

Suppose Pierre is a normal French speaker who lives in France and speaks not a word of English or any other language except French. Of course he has heard of that famous distant city, London (which he of course called ‘Londres’) though he himself has never left France. On the basis of what he has heard of London, he is inclined to think that it is pretty. So he says, in French, “Londres est jolie.”... Later, Pierre, through fortunate or unfortunate vicissitudes, moves to England, in fact to London itself, though to an unattractive part of the city with fairly uneducated inhabitants. He, like most of his neighbours, rarely leaves this part of the city... He learns, of course—speaking English—to call the city he lives in ‘London’. Pierre’s surroundings are, as I said, unattractive, and he is unimpressed with most of the rest of what he happens to see. So he is inclined to assent to the English sentence: [London is not pretty.]... Of course he does not for a moment withdraw his assent from the French sentence, “Londres est jolie”; he merely takes it for granted that the ugly city in which he is now stuck is distinct from the enchanting city he heard about in France. But he has no inclination to change his mind for a moment about the city he still calls ‘Londres’. (op. cit.: 254-255)

By means of the proper names ‘Londres’ and ‘London’, Pierre seems to have incompatible beliefs, though he is not irrational in doing so. In the same paper, Kripke supplies another case:

Peter... may learn the name ‘Paderewski’ with an identification of the person named as a famous pianist. Naturally, having learned this, Peter will assent to “Paderewski had musical talent,” and *we* can infer—using ‘Paderewski,’ as we usually do, to name the Polish musician and statesman: [Peter believes that Paderewski had musical talent.]... Later, in a different circle, Peter learns of someone called ‘Paderewski’ who was a Polish nationalist leader and Prime Minister. Peter is skeptical of the musical abilities of politicians. He concludes that probably two people, approximate contemporaries no doubt, were both named ‘Paderewski.’ Using ‘Paderewski’ as a name for the *statesman*, Peter assents to “Paderewski had no musical talent.” (op. cit.: 265)

This case seems to show that something like Pierre's situation can happen even with the same proper name in the same language.

Kripke intended his two cases purport to provide a puzzle that shows something about proper names that can be missed if one concentrates only on classic Frege cases. While the problem of referential coordination is a problem about thought, and not about language, the case of Peter and Paderewski/Paderewski does provide an interesting set of analogous cases for thought. One might describe Peter's psychology as thinking that Paderewski has no musical talent, and thinking that Paderewski has musical talent; and yet, Peter is not irrational in entertaining referentially inconsistent thoughts. Put in terms of ways of thinking, there is a question of how one should one distinguish ways of thinking in such cases: is the thinker employing the same way of thinking about Paderewski with both thoughts, though the thinker does not know this? If so, that would suggest something unexpected about ways of thinking, namely that they do not necessarily go with patterns of referential coordination. If not, then what is the basis for distinguishing them? The difficulty seems to be that there is a sense in which Peter should be using the same way of thinking, but fails to do so. This suggests that such cases need to be understood in terms of what explains referential coordination more generally.

A similar point can be made about some of the arguments made by Kripke (1980) concerning the issue of *descriptivism*. Descriptivism is the claim that reference to individuals goes by means of descriptions that those individuals satisfy, where a description is some sort of general condition that may be satisfied by distinct individuals in different possible ways for the world to be. Though directed against descriptivism about the semantics of proper names (what he called the 'Frege-Russell' view), Kripke's arguments, involving cases where it seems that descriptivism predicts the wrong results about reference and use, present a parallel set of issues about thought. The debate about descriptivism is an important one, though for present purposes, it can be overlooked. It seems clear that a descriptivist approach to thought about things will not present any distinctive way of dealing with referential coordination; if it does, it will only be under the auspices of one of the views of thought that I will consider in subsequent chapters.

The presentation of the cases provided by Perry and Evans, in contrast to Kripke's cases involving proper names, are described using different sorts of referring expression, those whose reference can change between given uses but whose meaning stays the same across those uses. In Kaplan's (1989) influential terminology, token expressions of this sort have

the same *character* as other token expressions of the same type, but can differ in their *content*. Indexicals, or ‘pure indexicals’, such as ‘I’, ‘now’, and ‘here’, and demonstratives, such as ‘this’ and ‘that’, fall into this category. Pure indexicals shift their reference as a function of elements of the context that can be specified independently of their being used, while demonstrative expressions shift their reference as a function of salience-making features of their use (paradigmatically, indicating by pointing at something). The thoughts in the Perry/Evans cases are often referred to as ‘indexical’ or ‘demonstrative’ thoughts, part of the reason being that they are most naturally expressed, and perhaps require expression, using these sorts of expressions.

Thoughts labelled as demonstrative are often held to have an indispensable explanatory role in tasks like engaging in actions that interact with objects in one’s immediate environment. Consequently, they are often closely tied to perceptual experience, or are said to be ‘perceptually based’ thoughts. Thoughts that are labelled as indexical are similarly held to have a crucial explanatory role in beliefs about ourselves, such as where we are, who we are, and what the present time is. Perry, in his discussion of the case (amongst several others) mentioned above, coined the phrase *essential indexicality*, and it is widely accepted that the thoughts in those cases are special in some way that answers to this description, and that there is a link between thoughts being essentially indexical (or demonstrative) and their connection to perception, identification, and action. The idea is that, just as there are expressions with the peculiar semantic structure, having a fixed character and variable content, so there are thoughts with an analogous representational structure, and this structure is somehow crucial for certain cognitive capacities.

The question of whether there really are thoughts of this kind is not independent of what should be said about referential coordination more generally. The existence and use of expressions with the character/content structure does not necessarily, or even obviously, bear on whether thoughts involve some analogous structure. That we can engage with objects in our immediate environment and think about ourselves is not an incidental feature of our cognitive lives. But it is not obvious that a special representational mechanism is needed to account for these capacities. Rather than the terminology of indexical and demonstrative thoughts, I will refer to this question more neutrally as *the problem of special ways of thinking*: is there a special representational structure that accounts for the rationality of identification and action on the basis of perception, and beliefs about ourselves?

Referential coordination may be treated as explanatorily prior to the problem of special ways of thinking. In order to determine whether special representation structures are needed, it needs to be seen whether an account of referential coordination that does not make use of any special representational structure can be made consistent with general facts about these capacities. If it can, then the motivation for thinking that there are indexical or demonstrative thoughts is to that extent undermined.

1.3.3 Coreference *de jure*

A reader familiar with the recent philosophical literature might have noticed a close affinity between referential coordination and what has been discussed under the heading of *coreference de jure*. This idea gets its most obvious application in the context of referring expressions that can be used to refer to the same thing previously referred to in a discourse, and need to be taken to do so in order to be properly understood. This covers the phenomena of syntactic binding, as with anaphoric pronouns, and repeated occurrences of noun phrases, and more complicated cases involving different sorts of referring expressions like proper names and definite descriptions. The expressions exhibit coreference *de jure* because of this condition on how they are to be understood. Coreference *de jure* is therefore to be contrasted with coreference *de facto*, which occurs when referring expressions refer to the same thing, but it is not a condition of understanding them that they are taken as doing so.

Although there is no definitive consensus on what coreference *de jure* is, we can see it as a combination of epistemic and semantic or representational conditions. Pinillos (2011: 303-5) cashes out coreference *de jure* in terms of three distinct epistemic conditions: a priority of existential generalisation (a relevant existential generalisation is known to follow *a priori*), attitude closure (the point about existential generalisation also goes when the sentences are embedded within propositional attitude contexts) and knowledge of conditional coreference (if two terms are coreferential *de jure*, then if they are understood, then they are known to refer to the same thing if they refer at all). Fine (2007) provides a similar set of conditions on what he calls *strict coreference*, or representation *as the same*. Schroeter (2012) discusses what she calls the *appearance of de jure* sameness of reference, and imposes the conditions of obviousness (no further reflection is required for one to figure out if they refer to the same thing), rational incontrovertibility (one cannot rationally reject sameness of reference if it is put to the

question), and epistemic primitiveness (there is no further thought to appeal to in explaining why one knows that they corefer).

Coreference *de jure* clearly cannot be assimilated to Frege cases or the puzzle about identity. It involves sameness rather than difference in ways of thinking. It may seem as though the terms used to articulate coreference *de jure* are very similar to those used to articulate referential coordination. This is particularly so in the idea that the one can infer an existential generalisation from things that are coreferential *de jure*. Nevertheless, they are not the same thing. Coreference *de jure* is much stronger than referential coordination. It involves a representational condition and an epistemic condition (or conditions), neither of which are part of referential coordination. The representational condition is coreference: two things corefer if and only if they refer to something and what they refer to is the same thing. The epistemic conditions involve someone knowing, or it being rationally incontrovertible, that the items (words, thoughts) are coreferential.

Referential coordination does not in itself involve coreference. Though it is natural to set up the problem by providing cases of thoughts that share referential content, to draw attention to the fact that referential content is insufficient, referential coordination does not require thoughts to refer to anything. Thoughts about things that do not exist can be referentially coordinated: Lucy is unwittingly made to participate in an experiment involving an hallucinogenic chemical and hallucinates a butterfly that is green and that it has wings, and so comes to think that there is such a butterfly; she can be rational in being disposed to infer that there is something that is a butterfly and has green wings, even though there is no butterfly and so her thoughts do not corefer because they do not refer at all. We could say that Lucy has a reason for treating her thoughts as coreferential *de facto*, though that would be unnecessarily prolix.

Recanati cashes out *de jure* sameness of reference in terms of knowledge of *conditional* coreference (2012: 92-3), which is knowledge that two things refer to the same thing if they refer at all, and so it might be thought that coreference as such is not required on some ways of understanding it. If so, then it remains the case that *de jure* coreference is stronger because referential coordination does not involve knowledge in this way. The paradigmatic inference is one that Lucy is disposed to make, and she is rational in doing so, but this does not necessarily mean that she knows that the paradigmatic inference follows, and she might not know anything about the inference; it just means that she has a reason to make it.

Moreover, granting that coreference *de jure* does involve sameness in ways of thinking, it must be said to do so in a way very different to referential coordination. Coreference *de jure* as a feature of language articulates a basic condition on linguistic competence that entails a certain level of reflectiveness. That is, it concerns the kinds of attitudes and evaluations we can take towards words and the sentences they compose that we need to have in order to be competent with a language, or to understand a given utterance or inscription. Similarly, as a feature of thought, the sorts of attitudes and evaluations that come with coreference *de jure* entail a level of reflectiveness about one's thoughts, and so a degree of cognitive sophistication, that is not entailed by referential coordination—even if, as Schroeter puts it, no *further* reflection is required, coreference *de jure* still seems to involve the capacity to reflect on one's thoughts.

The two conditions are therefore distinct, even though they both involve thoughts exhibiting sameness in ways of thinking. Referential coordination may also be treated as a prior explanatory problem, for two reasons. One is that referential coordination will obtain whenever coreference *de jure* obtains, though not vice versa. It is plausible to think that what explains the stronger condition should turn on what explains the weaker condition. Another reason is that coreference *de jure* may not, after all, be a genuine feature of thought. Although it is plausible to think that coreference *de jure* is a genuine aspect of the meaning of referring expressions, the combination of representational and epistemic conditions does not obviously apply to thought. That is not to say that it does not apply, only that whether or not it does is a substantive question that, presumably, will turn on the correctness or otherwise of theoretical accounts of thought. This in turn needs to be informed by the best explanation of referential coordination.

1.3.4 The metaphysics of intentionality

The fourth and final problem is not one that points to the insufficiency of referential content, but rather concerns the fact that thoughts have referential content at all. Following Brentano's (1874/1973) revival of Scholastic terminology, what is known as the metaphysics of intentionality concerns the conditions that are necessary and/or sufficient for one thing to be *about* another. Intentionality presents a host of problems. *Local* problems involve accounting for intentionality in particular cases: whether and how it possible to think or talk about what does not exist; whether there is a genuine distinction between general or descriptive aboutness (where one thing is about another in

virtue of it satisfying a certain collection of properties) and singular or particular aboutness (where one thing is about another independently of it satisfying any collection of properties); whether the capacity to be in states or undergo episodes with intentional content depends on the possession of inferential capacities, or whether they involve representational entities that have intentional content independently of inferential capacities.

These local problems have been discussed both in relation to and apart from the *global* problem concerning intentionality as such: as to whether it is a genuine property, and if so, what it is. One of the main debates on this score concerns the viability of a naturalistic theory of intentionality, such as those advocated by Dretske (1981), Millikan (1984), Papineau (1987), and Fodor (1990), on which intentionality is to be explained in non-intentional terms. Opposed to this are positions that accept the reality of intentionality but reject its explanation in naturalistic terms, for which Wittgenstein (1953) stands as the main progenitor. Opposed to both are eliminativist or instrumentalist (or in any case, not straightforwardly realist) attitudes towards the intentional, an approach that can be traced to Quine (1960).

The metaphysics of intentionality presents a problem that is in some sense deeper than referential coordination. But referential coordination may be treated as explanatorily prior in this case as well. Investigation into the conditions on intentionality requires having in hand a view of the kinds of things that exhibit it. Since one of the questions that gets raised by referential coordination is what view of thought needs to be adopted, settling what view provides the best explanation of referential coordination should be an input into the debates surrounding the metaphysics of intentionality. One cannot settle what the problem of the metaphysics of intentionality is really about, at least with respect to the intentionality of thought, without having a good idea of what the best explanation of referential coordination is.

A consequence of this is that explanations of referential coordination should not be evaluated in terms of or informed by considerations bearing on the viability or otherwise of *naturalism*, the view that every phenomena is either susceptible to explanation on the model of natural scientific theory, or not genuine. Rather, the viability of naturalism or its denial depends on the consequences of the right theory of referential coordination. In making this claim, I particularly mean to diverge from those, such as Fodor (1987; 1990; 2008), who would insist on a broad-ranging and thorough-going naturalism as a

presupposition of arguments concerning the right way to understand intentionality, and so also as a constraint on what the correct theory of mental content must be. Whatever exactly naturalism entails, it is a view that should be earned rather than claimed as a condition of any reasonable approach. In light of the fact that no proposal for naturalising intentionality commands anything like widespread acceptance (cf. Loewer 1997), I take it that doing things this way is more attractive than the Fodorian alternative.

1.3.5 Summary

Referential coordination presents an explanatory problem that is closely related to a range of other phenomena that have received comparatively much more attention. Referential coordination may be viewed as an independent problem in its own right. There are also good reasons to treat it as explanatorily prior to these problems. The correct explanation of referential coordination can provide a way of seeing what these problems involve and so how they might be resolved. The fact that referential coordination is implicit in these phenomena, and the fact that they have received a great deal of philosophical attention, attests to its importance.

1.4 *Questions and Theories*

This section sets out some questions to which possible explanations should provide answers, and introduces the explanatory theories to be discussed in subsequent chapters.

1.4.1 Three questions

Thoughts that are referentially coordinated can be thought of in two ways: in terms of their content, or in terms of the fact that they are things that have content. This provides a simple way to distinguish possible explanations of referential coordination. A *content-based* explanation looks to provide a theory of the contents of thoughts that allows for them to be more fine-grained than reference, and explains referential coordination by locating the indication of sameness of reference there. An opposed approach is instead to give a *vehicle-based* explanation. A vehicle-based explanation will look to provide a theory of the vehicles of thought, of the psychological states that have content, and explains referential coordination by locating the indication of sameness of reference in a fine-grained representational structure.

Though the contrast between content- and vehicle-based explanations provides a fairly natural way of thinking about possible positions and what is at stake between them, the contrast is not really adequate for a few reasons. It is not exclusive, in that there can be views that count as both. It is also inaccurate, in that there are distinctions that can be made within the two sides of the contrast that it leaves out. And it is incomplete, in that there are views that count as neither. More subtle distinctions can be made according to how views come down on three key questions.

One of the main questions concerning the explanation of referential coordination is:

Can referential content be treated as a primitive property, or not?

Call the view that it is *referentialism*, and the opposed view *anti-referentialism*. According to the referentialist, there are no facts more basic than thoughts' having the referential content they do that would bear on rational inference. The anti-referentialist, on the other hand, wants to appeal to the conditions that account for thoughts' having the referential contents that they do and use these conditions to provide a way explaining rational inference. As far as referentialism is concerned, thoughts may only have referential content (it is therefore to be distinguished from the stronger thesis that content is exhausted by reference), whereas the anti-referentialist is committed to there being some sort of content in addition to referential content.

A further question is:

Does the correct explanation of referential coordination make reference to psychological entities, properties, and states, or not?

Call the view that it does *psychologism*, and the opposed view *anti-psychologism*. The need for the psychologistic answer arises particularly acutely on the referentialist answer to the first question. If the explanation of referential coordination is not to be found in the conditions on referential content, then some other relevant feature must do the job, and psychological states and properties, of some sort, seem to be the only available alternative. Despite this, these first two questions are logically independent, in that combinations of views across both are perfectly coherent. One could advocate an psychologistic and anti-referentialist explanation as much as an anti-psychologistic and referentialist or

psychologistic and referentialist explanation. This demonstrates how thinking about the debate in terms of a contrast between content- and vehicle-based explanations is neither exclusive nor accurate.

A third question adds another dimension to possible kinds of explanation:

Is there an explanatory link between referential content and referential coordination, or not?

According to *heavyweight* theories, there is some explanatory link; according to *lightweight* theories, there is not. Lightweight theories must not be misconstrued as denying that there is any relationship at all between them. The point is that, on a lightweight theory, there is no *explanatory* link; what explains one need not figure in an explanation of the other. As with the first and second questions, there is some overlap between answers to this third question and the previous two. Lightweight explanations of referential coordination must be referentialist, and so must be psychologistic. But they will be neutral as to whether the psychological states or properties must also be states or properties that involve referential content. This neutrality is what makes them lightweight; they are relatively non-committal as to what needs to be in place. If properties relevant to determining referential content play no part of the explanation of referential coordination as on the lightweight approach, then it is wrong to say that it is either content- or vehicle-based explanation either. This shows that the contrast between content- and vehicle-based explanations is also incomplete, because there can be explanations that make use of neither notion.

1.4.2 Theoretical approaches

Explanations of referential coordination can be seen as implicit in some theoretical approaches to thought that have a prominent place in the philosophical literature. The three questions go a long way towards characterising both what is at issue between them and how they overlap. The division between the approaches is in some ways artificial, but for purposes of orientation it helps to see them as discontinuous competitors.

Going back to the initial crude distinction, we can begin with theoretical approaches that offer explanations in terms of contents and vehicles.

According to *sense theory*, thoughts have contents additional to their referential contents. These contents are structured complexes the constituents of which are individuated by their role in rational inference, and that determine referential content in a fine-grained way. These constituents therefore have properties that make them akin to Frege's notion of sense. Sameness of sense indicates sameness of reference, because the same referential content is determined in the same way. Sense theory represents one development of Frege's notion of sense, introduced by him in the course of dealing with some problems in semantics, which has not proved uncontroversial even amongst supporters of Frege.

Sense theory has often been opposed by defenders of *typed representations* theory, on which thought involves mental representations. These are vehicles of referential content that have causal roles in a cognitive system, structured in a way akin to how a sentence is made up of words. The types of typed representations theory are ways of grouping such representations together. The representations grouped together are tokens; the groupings are the types of those tokens. Types of mental representations are meant to involve properties that bear on the way that the thinker whose mental representations they are can manipulate them in cognitive processes. The thought is that properties that type these representations can provide a similar fine-grained relationship to referential content without the need to bring in an additional notion of content. On this approach, referentially coordinated thoughts are a matter of mental representations belonging to the same type.

Both approaches are internally diverse, and it will be necessary to spend some time articulating the details of how they work in due course. At this stage we can remain at the level of the general ideas upon which each approach relies. They both offer explanations that can be characterised as heavyweight, in that they locate the explanation in one or the other side of the content/vehicle contrast. While sense theory is explicitly anti-referentialist and anti-psychologistic, typed representations theory is explicitly psychologistic, and depending on the typing criteria, can be either referentialist or not. Typed representations have often been appealed to in light of various difficulties that have been alleged against the anti-referentialist features of sense theory, and the sorts of explanation that I will consider under that heading will be those with referentialist commitments (see §4.2).

A third explanatory tradition that can be distinguished from both sense theory and typed representations theory is *mental file theory*. The basic idea of mental file theory is to posit

mental entities that are like the sorts of files one finds in draws and cabinets, in that they are associated with, or in some sense contain, information, and organise that information in useful ways. Unlike senses, though like typed representations, mental file explanations are psychologistic. Unlike both, they need not, though they can, be involved in explanations of referential content. Mental file explanations can therefore be either heavyweight or lightweight, and either referentialist or anti-referentialist. Extant treatments of mental file theory tend to be given in the service of heavyweight approaches to thought.

The three theoretical approaches obviously overlap and differ in many and more-or-less complicated ways, but one can see that between them, and taking into account internal disagreements, they cover most of the positions defined by the three questions. The exception to this is a position that is anti-psychologistic and referentialist. It is hard to find examples of this sort of view in the literature, perhaps because it is difficult to see how fine-grained contents that are not individuated in terms of rational inferences can bear on a thinker's inferential dispositions at all. Fine's (2007) proposal to introduce equivalence relations, or *coordination schemes*, over referentialist propositions as part of the contents of thoughts may be an example, although it is unclear if it represents a genuine explanatory alternative to the three views just sketched (cf. Lawlor 2010, Sosa 2010). In any case, I will overlook this alternative as insufficiently developed.

These theories and their respective positions on the three questions can be set out in table form like so:

	Psychologism	Referentialism	Weight
Senses	No	No	Heavy
Typed representations	Yes	Yes	Heavy
Mental files	Yes	Yes/No	Heavy

As this table indicates, there is logical space for a position that differs from the other three in at least one respect. This space is occupied by the explanation I will develop and defend.

1.4.3 Introducing coordination functions

I will develop and defend an explanation that uses mental file theory, extended in a couple of ways. One extension is a distinction between mental files and file *predications*. Files and their predications provide the basis for the fact that a thinker has the paradigmatic inferential dispositions (note: not their rationality). The second extension is the idea that the role these dispositions have in a thinker’s life, in structuring and providing information relied upon in furthering her broader projects and plans, mean that file predications have normative functional properties that concern the information that they carry. File predications belonging to the same file are *required* to carry information about the same thing. This coreference requirement means that file predications belonging to the same file have *coordination functions* that link them. Requirement-based coordination functions put in place a default indication that the information carried in is about the same thing. Information that indicates that the conclusion cannot or is unlikely to be true will defeat this. A special kind of defeating information will concern the identity of the thing that the information concerns. The rationality of the paradigmatic inferences consists in a thinker’s sensitivity to the coherence of her inferential dispositions with the rest of her information and the potential for defeat that it presents.

This explanation fits into the framework like this: it is *psychologistic* because mental files and their predications are mental entities; it is *referentialist* because it makes no use of and is not dependent on any more fine-grained notion of content; and it is *lightweight* because there is no explanatory link between this explanation and the explanation of the referential content of thoughts.

Adding coordination functions to the table given above makes it look like this:

	Psychologism	Referentialism	Weight
Senses	No	No	Heavy
Typed representations	Yes	Yes	Heavy
Mental files	Yes	Yes/No	Heavy
Coordination functions	Yes	Yes	Light

The character of the account I wish to propose is, then, somewhat unique. One may think that there are good reasons for this. Something that I acknowledge from the start is that it is not especially *inspiring*. In particular, it does not do much by way of furthering certain projects in epistemology, philosophy of mind, or philosophy of language. This point comes out quite clearly by noting how theories in these domains, and accounts of inference, and thought in general, are often given in terms of *concepts*. Concepts are often made use of in the course of pursuing broader philosophical projects. One example of such a concept-theoretic approach is the attempt to provide an account of a special sort of knowledge, often called *a priori*, that someone can have independently of experience. This sort of rationalist epistemology is sometimes supported by the idea that the conditions needed for possessing certain concepts also supply epistemic warrant of a distinctive kind. Another example is the attempt, consonant with the project of cognitive science, to provide a theory of the capacities that thinkers have to be in mental states with referential content, and the systematic relationships between them, compatible with the fact that we are creatures with a finite cognitive repertoire. Referential coordination by itself speaks to neither of these issues, and this may seem a disappointment and reason enough to dismiss it.

Those who are sympathetic to that thought are asked for forbearance, however. Though the uninspiring—or, as I would prefer to put it, relatively deflationary—explanation that I want to recommend does not speak directly to ways of providing a rationalist epistemology, or furthering the project of cognitive science, a benefit of focusing on referential coordination is that it provides a sort of key for the various uses (and there are many) to which talk of concepts is put. This means that they can be comprehended as at least bearing on the same topic, and so apparently different explanatory concerns can be made to bear on each other in a philosophically satisfying way. This point will be discussed in greater detail in the next chapter.

The result of focusing on referential coordination is not limited to offering a way of clearing things up, however. A commitment to explaining features of thought in terms of concepts is a commitment to a heavyweight explanation of referential coordination, since concepts are meant to explain both the referential properties and the rational properties of thoughts. If an explanation that only requires a lightweight relationship between referential coordination and referential content is correct, then no support for a rationalist epistemology or an account of referential content is provided. That these projects cannot be supported through reflection on referential coordination, and that the

soundness of talk of concepts is questionable, are interesting results even though they have deflationary consequences. Those with more inflationary ambitions owe a response other than hopeful dismissal on this score.

1.4.4 Summary

Different theories of thought can be seen as differing over how they explain referential coordination. Different explanations differ over what provides the inferentially basic indication of sameness of reference sufficient for making the paradigmatic inferential dispositions rational: whether it involves a more fine-grained notion of content, whether it involves psychological entities and properties or not, and whether it is linked to the possession of referential content or not. The coordination functions explanation that I will defend provides a distinctive set of answers to these questions.

1.5 *Assumptions, Requirements, and Desiderata*

I will now set out the explanatory criteria that will be applied to the solutions to that problem to be considered in the remaining chapters. This has three parts: first, setting out some assumptions that are in place that help to focus the scope of the explanations; second, setting out some requirements on explanations that must be met as a standard of adequacy; and third, setting out some desiderata on explanations that meet that minimal standard.

I will rely on the following assumptions:

Core Cases: The explanation of referential coordination should, in the first instance, be given in terms of cases where a thinker has thoughts about individuals in a straightforward sense, or *core cases*.

Core cases are cases of propositional attitudes that do not introduce any more delicate problems than that of referential coordination in order to understand what is going on. Peripheral cases include thoughts about ourselves, our present location and present time, the kinds of thoughts covered by what I earlier called special ways of thinking. Peripheral cases also include cases where a thinker is arguably exploiting some description, cases of thoughts about fictional or non-existent entities, cases of thoughts that rely on some kind

of deference to experts, or any logical notions more sophisticated than those engaged in the paradigmatic inferences.

Availability: In core cases, referential content is unproblematically available, in the sense that what we think about, and that we think about them, is easy to specify.

Cases such as Lucy's thinking about Cicero are in the core; it is reasonably clear what is going on, and her thinking about Cicero, in her thinking that he is a Roman, does not obviously involve any peculiarities that demand particular treatments of her thoughts. Rather than start with oddities and work back to the more mundane cases, these assumptions mandate dealing with the more ordinary cases and then building outward, if needed. In the core cases, the fact that someone is thinking about something does not itself raise any puzzles, and so can be treated as a given for the explanation. Any reasons to think otherwise will require assumptions that are more controversial than this, and so can be put aside for the purposes of evaluating explanations of referential coordination.

An explanation of referential coordination must meet these three conditions of adequacy:

An explanation must specify an indication of sameness of reference that:

- (1) provides a reason for the thinker to engage in the paradigmatic inferences that is
- (2) inferentially basic, and
- (3) adequately personal-level.

Requirement (1) means that it is not enough to come up with some feature of thoughts that provides an indication of reference that is only available from the theorist's perspective; it must be something that the thinker can respond to. Requirement (2) means that the indication should not consist in something that needs to be included as a premise in the paradigmatic inferences. Requirement (3) means this inferentially basic reason cannot operate only at the sub-personal level. An explanation must join up with the idea that it makes sense for a thinker to answer the question as to why they are disposed to make the paradigmatic inference by saying something to effect of "what I was thinking of is one and the same thing." The target phenomenon is one that concerns a thinker (a person) and her perspective on her thoughts, how they make sense to her. Any

explanation that does not link up with what a thinker can be conscious of as expressed in this sort of answer is inadequately personal-level.

However, we need not find ourselves, in Martin's (1997) phrase, "stuck here in the shallows of the mind." A personal level explanation, in the sense that figures in the requirement, is not one that can *only* make reference to that of which a thinker can be conscious. An explanation needs to make contact (as we might put it) with the personal level through showing *why* it would make sense for a thinker suitably equipped with the verbal and reflective resources to cite the identity of what she thinks about. But making contact in that way does not limit the resources that can be legitimately used in giving an explanation to those found in the conscious life of a thinker. If a psychologicistic explanation is correct, then something beyond this will be required, and there is no sense in ruling out psychologicistic explanations on the basis of an arbitrary restriction.

I will impose the following desiderata on explanations that meet the conditions of adequacy:

The explanation to be preferred provides a specification that is:

- (1) sufficiently general,
- (2) non-circular, and
- (3) psychologically realistic, in a way that
- (4) can be extended to providing a plausible account of identity judgements, and
- (5) is consistent with general facts concerning special ways of thinking.

Desiderata (1-3) are applications of general explanatory virtues to the specific case. They are meant to apply to the core cases, as per the assumptions noted above. (1) means that an explanation is better if it covers these cases; (2) means that an explanation is better if it does not make appeal to or introduce some explanatory notion that can only be justified on the basis that the theory is correct, or that needs to itself be explained in terms of the capacity to have referentially coordinated thoughts; (3) means that an explanation is better if it does not require of thinkers capacities or other properties that we have no reason, independent of the explanatory proposal, to think they have, or that are overly demanding in proportion to what it generally takes to be a thinker capable of referentially coordinated thoughts.

Desiderata (4) and (5) are an application of the idea that the explanation of referential coordination is not wholly independent of closely related problems discussed above. Identity judgements overcome differences in ways of thinking. Special ways of thinking involve referentially coordinated thoughts that issue in particular cognitive abilities, such as intentional action or thought about oneself and one's present location. It is, therefore, incumbent on a theory that it be consistent with general facts about these cognitive abilities and does not render them mysterious or unavailable. I emphasise that it must be consistent with general facts, not with claims that can only be motivated or articulated on theories concerning the nature of thought that are open to disconfirmation by dint of how they fare as an explanation of the core cases as per the other desiderata.

The reasons for adopting the coordination functions explanation can be briefly stated as follows. Anti-psychologistic explanations, as represented by sense theory, and psychologistic explanations as represented by typed representations, are both involved in problems to do with circularity and psychological realism. Extant treatments of mental file theory suffer from being psychologically unrealistic, or else from failing to cover enough cases. These problems can be avoided by adopting lightweight approach, which in turn means adopting referentialism.

1.6 *Conclusion*

Referential coordination is a way for identity to figure in thought that is distinctive in that it involves an inferentially basic though personal-level indication of sameness of reference. The problem is that it is not obvious what this indication might be. The problem of referential coordination is closely related to some other problems that have been relatively more widely discussed, but it should be treated as independent and explanatorily prior to these. There are a number of theories that differ over how they address some key questions about the nature of thought, on which the theory I will defend offers a distinctive set of answers. Arguing for it requires demonstrating that it meets the requirements and desiderata on an explanation better than the competitors.

Chapter 2 - Concepts

2.1 Introduction

This chapter discusses the relationship between the problem of referential coordination and the theory of concepts. Readers impatient to get to the main argument of the essay may skip this chapter without much loss, although it contains material germane to how that argument relates to the philosophical literature.

There are a number of views on what concepts are, and these are different enough to encourage the thought that they are not really views on the same thing (§2.2). I propose that these competing view may be understood as diverging over the nature of theoretical entities that explain why thoughts that are exercises of referential capacities are also exercises of what I term *conceptual capacities*. A conceptual capacity is exercised in thoughts that are referentially coordinated. The link between referential and conceptual capacities means that concepts are tied to a heavyweight explanation of referential coordination (§2.3). This does better than some alternatives at comprehending what is at issue between competing views of concepts (§2.4). The term ‘concept’ is also used in psychological theories of categorisation and related epistemic capacities, and this provides a contrast between psychological and philosophical concept theory; I argue that whether and how these are related depends on what the best explanation of referential coordination is (§2.5). More generally, the cogency of (philosophical) concept theory depends on the merits of a heavyweight explanation of referential coordination; if the argument in subsequent chapters is correct, then a lightweight theory is better and the commitment to concepts will to that extent be undermined (§2.6).

2.2 What Are Concepts?

It is a banal point that *concepts* are involved in *conceiving of things*, which is to say, in *thinking* about them. But in what way? What are concepts? This is not an entirely straightforward question to answer.

The fact that ‘concept’ is a term that gets used to refer different things is often remarked upon (see, for example, Peacocke 1992: 1-2). Some of these are highly distinct, some more closely related. Margolis and Laurence (1999), in a lengthy survey of the literature on concepts, find them variously identified with prototypes, exemplars, theories, various

hybrids of these things, definitions both classical and neo-classical, and unstructured symbolic atoms. In other parts of the literature, one finds concepts identified with mental images, or abilities to use words, or with mental files. One might well wonder what it is that makes putting all of these things together under the heading of concepts appropriate.

With some qualification, we can distinguish two ways in which the question ‘what is a concept?’ might be taken. On one way of taking it, the question is about what it is that concepts *do*, or what their *explanatory role* is; call this the *job description* question. On a different way of taking it, the question is about what sort of thing it is that *plays* this role; call this the *identification* question. The identification question asks what sort of thing, otherwise picked out, should be identified as that which plays the concept role. A great deal of the difficulty with making sense of all the many different uses of the same term arise from the fact that these questions are not as easy to keep apart as they are to distinguish. The way in which one is answered will constrain and inform the answer to the other.

This problem can be illustrated by considering a way that Margolis and Laurence frame their survey of the concepts literature (1999: 5-8, 75-77). As they say, concepts are often identified at a high level of abstraction as being *mental representations* of categories, properties, individuals, or whatever else. As such, concepts are sometimes thought to be the *vehicles* of mental content, or otherwise of various kinds of information. This is contrasted with an opposed identification of concepts as parts of the *contents* of mental states. This seems like a straightforward disagreement (recall the discussion in §1.4). But do these identifications answer to the same job description? If not, then are they really competitors?

In line with this divergence, debates about concepts are sometimes taken to be concerned with whether concepts are to be thought of as abstract or non-abstract entities (see Margolis & Laurence 2007; Sutton 2004). While this may well be something at stake between proponents of the different answers to the identification question, it is unlikely to be what the disagreement is about at a more fundamental level. The fact that concepts as constituents of contents are abstracta is less important than the fact that the claim that they are constituents of contents is made, presumably, in the course of propounding a particular theory to a particular explanatory end. As a representative example, Burge (1993a: 309) claims that a benefit of the claim that mental contents are conceptually

structured (i.e. have concepts as constituents) is that such structures “enable one to capture a thinker's ability to relate different thoughts to one another according to rational inferential patterns.” What drives the identification of concepts with abstracta is this explanatory concern, not some ancillary ontological agenda. By contrast, the identification of concepts as vehicles of mental content is more focused on a different collection of explanatory concerns. Again, as a representative example, Fodor's identification of concepts with vehicles of thought has been directed at accounting for the systematic nature of intentional states and offering a naturalistic theory of them (see especially Fodor 2008: 25-49; also Fodor 1998: 1-22).

At the same time, it would be wrong to think that there is no overlap in broader explanatory projects. For instance, Peacocke (1992: 41-59), after explicitly identifying concepts as constituents of contents, offers a lengthy discussion of the systematic properties of thought, in the guise of a discussion of Evans' (1982: 100-105) concern with what he called the Generality Constraint. And going in the other direction, though Fodor (2008) makes clear that his concern with concepts is part of a project of a theory of the workings of the representational mechanisms of the cognitive system, he has a lengthy discussion of the way in which his favoured account deals with Frege cases of the various sorts discussed in the previous chapter (op. cit.: 50-100).

Both explanatory projects bear on the nature of thought, so it is not surprising that there should be overlap. But beyond this, more needs to be said to relate them, and from the perspective of figuring out whether divergent answers to the job description and identification questions can be comprehended as bearing on the same topic, the overlap is less important than the difference in fundamental explanatory concerns, which is big enough to give one pause. Different takes on the identification question can be informed by different takes on the job description question, and *vice versa*. The problem is that there is no obvious position from which to decide between them, and perhaps no need to decide between them, if this is a merely verbal disagreement.

To see the problem from another angle, consider what is a fairly common way of introducing talk about concepts, the dictum that *concepts are building blocks of thoughts*. One finds this given a variety of formulations, but the basic idea is the same: concepts are fundamental in making up thoughts, in some sense of ‘making up’. The problems that we saw above with the job description/identification questions arise for the dictum, as one might expect: what exactly is the sense in which they make up thoughts? and what is

meant by ‘thoughts’? Is it the content of a thought, or vehicle of that content? What it means to be a constituent of a thought depends on how this is answered, so the difference between content and vehicle provides for incompatible ways of understanding the dictum. While it perhaps serves to focus the issue, the dictum itself sheds little light.

Margolis and Laurence remark that “[g]iven the foundational role that concepts have for understanding the nature of cognition, it’s not possible to provide a theory of concepts without taking sides on a number of fundamental questions about the mind.” (2003: 190) Perhaps this is right, and perhaps it extends also to it not being possible even to say what a theory of concepts is a theory of without taking such a stand. It might then seem attractive, or even obligatory, to simply side-step the job description question altogether, and instead pick an answer to the identification question and stipulate that that is what concepts are. Peacocke (1992: 2-3) does something along these lines when he makes clear that his talk of concepts is intended to pick out constituents of propositional contents at the level of Fregean sense, and then goes on to discuss a wide range of issues related to thought, perception, and language. This sort of stipulation has the not negligible attraction of promising progress without getting bogged down in ground-level disagreements over fundamental questions about the mind. But it does mean that no stance is taken on those fundamental questions, and there is the danger that it makes it more rather than less difficult to avoid merely verbal disagreement.

Alternatively, and in the interests of *not* simply taking sides, one might think it a good idea to dismiss concepts as “creatures of darkness” (Stalnaker 2008: 105-6) and simply try to avoid talking about them at all. The problem with this is that the sorts of explanatory proposals framed in terms of concepts, and indeed much of what has been done in the direction of giving a philosophical account of thought, would be rendered opaque.

It would be far preferable to come up with a way of answering the job description question that does maximum justice to divergent positions on the identification question and involves as little side-taking as possible. Fortunately, there *is* such an answer that achieves this. As may come as no surprise, I want to suggest that the job description question should be answered in terms of referential coordination. Put in terms of the dictum: concepts are the building blocks of thoughts because they explain a fundamental fact about how thoughts work, that they exhibit patterns of referential coordination. This is as neutral as one could hope, and provides a reasonably clean way to separate out what

different ways of talking about concepts are about and to discern whether they are competing or consistent.

2.3 *Conceptual Capacities*

To see how referential coordination helps, rather than focus on thoughts just as such, which seems to force the question of whether we have contents or vehicles in mind, we might do better to look at them as exercises of certain capacities.

Thoughts can be viewed as exercises of capacities. Which capacities? There are a number of different ways that thoughts might be grouped together, and hence a number of different ways of individuating these capacities, where individuating such a capacity is giving sufficient conditions for a thought to be an exercise of it. One of these is in terms of what thoughts are about. We can call capacities individuated in this way *referential capacities*. Exercises of the same referential capacity are about the same thing. However, if exercises of referential capacities—that is, thoughts about something—are exercises of capacities of *rational* thinkers, then a further condition is required. For the purposes of enumerating the capacities of rational thinkers, thoughts need to be grouped together according to whether the thinker has a reason to treat them *as* being about the same thing or not. In other words, the thoughts of rational thinkers exhibit patterns of sameness and difference in ways of thinking. This means that the cognitive capacities exercised by rational thinkers are individuated not just by their referential content but also by their being referentially coordinated. We can call the capacities individuated this way *conceptual capacities*.

To illustrate: Lucy knows Fido the dog, and thinks that Fido has fleas. Thinking this is exercising a referential capacity, the capacity to think about Fido. Lucy also thinks that Fido is a labrador. In thinking this, she exercises the same referential capacity again. She is also disposed to infer, on the basis of these two thoughts, that there is something that is a labrador that has fleas, and it makes sense to her to do so. These two thoughts about Fido are therefore referentially coordinated. So these thoughts involve exercising the same conceptual capacity as well as the same referential capacity. Using a standard typographical convention, we can put this by saying that Lucy is exercising her FIDO-capacity in both cases. Lucy sees Fido shortly after a trip to the vet, but his new anti-flea collar has somehow been mixed up with one belonging to a dog called Rex. Lucy fails to recognise him and thinks of him as a different dog with that name. Knowing that the

dog she thinks of as Rex has an anti-flea collar, she thinks that Rex does not have fleas. In thinking this, she is thinking of Fido, and so is exercising the same referential capacity. But she is not exercising her FIDO-capacity. Perhaps she still thinks that Fido has fleas. She would not be irrational in maintaining this belief. Instead, she is exercising a different conceptual capacity, her REX-capacity. She also thinks that Rex is a labrador, and is rationally disposed to infer that there is a labrador that does not have fleas. These two thoughts are therefore both exercises of her REX-capacity.

Exercises of conceptual capacities are therefore exercises of referential capacities that meet some further condition that makes them referentially coordinated. We can then frame an interesting explanatory hypothesis about conceptual capacities: the fact that thoughts are exercises of the same conceptual capacity because they are constituted, at least in part, by the same entity. We could then call the entities posited by this explanation *concepts*. What makes the first two of Lucy's thoughts about Fido exercises of her FIDO-capacity are that they are partially constituted by the same concept, the concept FIDO. FIDO is absent in the third thought that is nevertheless about Fido. Instead, in exercising her REX-capacity, she thinks a thought constituted by a distinct concept, the concept REX.

The job description question can therefore be answered like this: concepts are whatever it is the presence of which explains why thoughts that are exercises of the same referential capacity are also exercises of the same conceptual capacity. Same concept, same referential capacity, same conceptual capacity; call this *the capacities answer* to the job description question.

What I will *concept theory* is the idea that there are concepts, that is, things that answer to this job description. The debate between those who hold divergent views on what concepts are can be seen as coming down to which answer to the identification question should be given. This is determined by which best solves the problem of referential coordination. The immediate benefit of this interpretation is that the problem of referential coordination can be posed *without* presupposing much by way of fundamental questions about the mind; we can seek a satisfying explanation of referential coordination without having to settle the other questions.

This helps make sense of much of the debate about concepts, and in particular with comprehending how different answers to the identification question are in fact related to the same fundamental explanatory concern. The emphasis on rational inferential relations

is an emphasis on what it is that explains the rational inferential relations that make them exercises of the same conceptual capacity. The emphasis on systematicity is an emphasis on how to explain this as a feature of conceptual capacities. These are different explanatory projects, but they are still parts of one overarching explanatory project, and answerable to a central strand of that project, the explanation of referential coordination.

Earlier, I said that taking the question ‘what are concepts?’ as either asking for a job description or for an identification of a kind of entity otherwise described needs qualification. The qualification is this: not all talk of concepts is intended to be explanatory. Reference to concepts can be a means of capturing what thoughts a thinker can have, either by simply adverting to what she can think about described without any reference to what identity relations her thinking presupposes, or else as a way of capturing these presuppositions. So, for example, we might say that Lucy has a concept of Fido (she ‘has the concept FIDO’), because she can think about Fido, and also has the concept of having fleas, because she can think of things’ having fleas. But she also has a distinct concept (she ‘has the concept REX’) because she can think of Fido in two different ways. This sort of reference to concepts, what we might call concepts in the *descriptive* sense, contrasts with the theoretical sense corresponding to the capacities answer.

The capacities answer has its uses here too. The descriptive sense brings with it no weighty theoretical presuppositions, whereas the theoretical sense does. This means that cases where the use is theoretical and where it is not should be held apart, and the capacities answer provides a means of telling which is which—and also for revealing cases where the distinction is elided (perhaps unintentionally) and some substantive commitment gets introduced under cover.

The capacities answer converges in an interesting way with Evans’ (1982) talk of *conceptual abilities* in the context of his discussion of the Generality Constraint (op. cit.: 101-5) mentioned above. The Generality Constraint is sometimes taken to be definitional of conceptual thought (see Beck 2012 for a recent example). Seeing thoughts as exercises of conceptual capacities is in line with this. Evans states the Generality Constraint like so:

[I]f a subject can be credited with thought that *a* is *F*, then he must have the conceptual resources for entertaining the thought that *a* is *G*. (1982: 104)

The idea that concepts are entities that explain how thoughts are exercises of the same referential and conceptual capacity involves no commitment to the kind of systematic connection between the thoughts a thinker is able to have to which Evans wants to draw attention. However, systematic connections aside, both the Generality Constraint and the explanation of conceptual capacities concern the fact that what thoughts a thinker can have are not arbitrary; there are connections between them, and these involve seeing thoughts as in some sense structured, and this structure corresponds to the capacities that are exercised in thinking a thought. This is made clear when Evans goes on to elaborate by saying that

someone who thinks that John is happy and that John is sad exercises on two occasions a *single ability*, the ability to think of, or think about, John. (op. cit.: 101)

The convergence becomes more obvious when, in a footnote, Evans adds

When two thought-episodes depend on the *same* ability to think of something, we can say that the thing is thought about *in the same way*. (op. cit.: 101n16)

The abilities that figure in the Generality Constraint are articulated not just in terms of referential coordination, but in a more fine-grained way. If one thinks of concepts as governed by the Generality Constraint, then one should also think of them as tied to the explanation of conceptual capacities.

Evans denies that mental representations are required by the Generality Constraint (op. cit.: 101), and instead recommends that it be articulated just in terms of abilities. The explanation of the Generality Constraint that he sketches and goes on to develop is not, however, neutral, but is clearly aligned with a content-based view of concepts, as is made clear in the comparison with Frege's notion of sense in the footnote just quoted (see also op. cit.: 104n24). One need not follow Evans on this. Concepts, understood as entities that explain referential coordination, might be more or less anything, and might not have

anything to do with Frege's notion of sense. The identification question remains open when the job description question is answered in this way.

On a related point, while adopting the capacities answer means that concepts are closely tied to the inferences a thinker can make, it does not beg any questions against those, such as Fodor, who deny that concepts are individuated by their role in inference, the position he terms 'Cartesianism' and contrasts with 'Pragmatism' (see Fodor 2004, 2008). The capacities answer does not mean that particular concepts play particular roles in inference, much less that what makes a concept the concept that it is has a particular inferential role. It only means that concepts play a role in a particular *form* of inference; using the same concept means being disposed to treat one's thoughts as being about the same thing, and this means being disposed to make inferences of certain forms, paradigmatically the inference to conjoin and generalise. This does not individuate concepts, it just sets a condition on what concepts, however individuated, do. If Cartesians and Pragmatists about concepts cannot agree even on this point, then it becomes unclear what the substantive issue between them might be.

2.4 *Alternative Answers*

The literature on concepts features other ways of answering the job description so as to comprehend different answers to the identification question. They do not do quite so well as the capacities answer.

One of these alternatives is the *requirements answer*. This sees concepts as the target of requirements that any adequate theory must respect; concepts should be identified with whatever respects them best. This sort of answer is advocated by Fodor (1998: 22-39), who places a number of 'non-negotiable' requirements on a theory of concepts. On his view, they must be

- (i) mental particulars,
- (ii) categories and routinely employed as such,
- (iii) constituents of thoughts (the dictum again), and so
- (iv) have compositional structure,
- (v) such that many are learned, and
- (vi) public, which is to say, sharable across different thinkers.

Although a little more sophisticated than simply offering a stipulation, this does not mark much of a departure. Someone *could* dispute any given requirement, and there would not a lot to fall back on other than agreeing to disagree. *Why* the requirements are requirements on a theory of concepts is left unclear. They must come from somewhere, and that somewhere must be some explanatory project. Simply imposing the requirements does little by way of helping with uncovering what this might be.

Perhaps a more promising alternative on this score is the *desiderata answer*. Prinz (2002) presents something akin to Fodor's list of non-negotiable demands, though in the form of explanatory desiderata to which he thinks a theory of concepts should be responsive, and which he finds "widely accepted among philosophers and psychologists." (op. cit.: 3) Rather than non-negotiable, they are negotiable provided some trade-off can be made between competing explanatory proposals. The explanatory project is cast as a collection of different explanatory projects. A theory of concepts is better if fewer trade-offs are required, and the more compelling the explanation given on each of the particular desiderata. The desiderata Prinz provides are

- (i) Scope: a theory should "accommodate the large variety of concepts that we are capable of possessing;" (ibid.)
- (ii) Intentional content: a theory should "help us understand how concepts attain their intentional content," (op. cit.: 4) i.e. their referential properties;
- (iii) Cognitive content: a theory should explain "how coreferential concepts can differ and how divergently referential concepts can be alike;" (op. cit.: 8)
- (iv) Acquisition: a theory should "ultimately support a plausible explanation of how concepts are acquired;" (ibid.)
- (v) Categorisation: a theory should predict and explain facts to do with the cognitive ability of categorisation (op. cit.: 9-12);
- (vi) Compositionality: a theory should account for the fact that the intentional and cognitive contents of complex concepts are a function of the simple concepts out of which they are composed. so as to account for the ability to form novel thoughts (op. cit.: 13);
- (vii) Publicity: a theory should account for the fact that concepts are "capable of being shared by different individuals and by one individual at different times." (op. cit. 14)

This list, based on the various uses to which concepts have been put, brings out the fact that concepts are often seen as explaining *more* than just exercises of conceptual capacities, in my sense. The capacities that concepts are supposed to explain include not just the fact that their exercises are coordinated (or come in systematic groups), but also that they include referential capacities, and are tied up with the exercise of other cognitive capacities such as identification and categorisation. Coming at the job description question with a list of desiderata may seem to provide a degree of flexibility not available otherwise.

The problem with the desiderata answer is that it provides no way to settle disputes over how the desiderata are understood, whether or how they should be ordered for the purposes of trade-offs, and whether they genuinely apply or not. To take an example of a disputed desideratum from Prinz's own discussion, he imposes the categorisation requirement because, although "[i]f concepts are thought constituents, it might be best to remove the explanation of categorisation from the list of desiderata," doing otherwise

would strongly bias the case against psychological theories of concepts. In psychology, an enormous amount of research on concepts has focused on categorisation. Concepts are often stipulated to be the cognitive mechanisms by which we categorise. If a theory of concepts were absolved of its obligation to explain categorisation, most psychological accounts would be rendered moot. (op. cit.: 11)

Prinz is surely right that "[c]ategorisation certainly stands in need of an explanation," (ibid.) but part of what is at issue in the debate about what concepts are is whether one and the same thing explains categorisation and plays the explanatory roles associated with conceptual constituents of thought contents, and so whether the obligation (presumably Prinz means the desideratum) can be imposed. What is needed is a way to decide this. That psychological accounts might be rendered moot is beside the point (more on psychological theories of concepts shortly).

While it is true that explanatory power achieved through the unified explanation of many problems is a virtue of a theory, it is not enough to come at the job description question without a principled way of comprehending the problems and the relations between them. Any way of unifying the less basic explanatory problems will presuppose some idea of what the best explanation of referential coordination is. Referential coordination

provides the principled way of comprehending the problems that is missing from the desiderata answer. So whatever speaks in favour of the desiderata answer speaks more strongly in favour of the capacities answer.

2.5 *Philosophical vs. Psychological Concept Theory*

What I have been calling concept theory is the idea that there are entities that explain why different thoughts that are exercises of the same referential capacity are also exercises of the same conceptual capacity. Following Machery (2009), we might locate this in a further distinction between *psychological* and *philosophical* concept theory. As Machery has it, concepts as understood by psychologists are “those bodies of knowledge that are stored in long-term memory and that are used in the processes underlying the higher cognitive competences,” (op. cit.: 10) where knowledge is understood in an epistemically neutral way, and the higher cognitive competences include categorisation, inductive and deductive inference, analogy formation, linguistic understanding, and planning. By contrast, Machery takes it that concepts as understood by philosophers are “capacities for having propositional attitudes.” (op. cit.: 31) This difference between psychological and philosophical concepts is most helpfully seen as a divergence over the job description question. Philosophical concept theory bears directly on the explanation of conceptual capacities; psychological theory bears directly on the explanation of behaviours linked to categorisation.

Prominent entries in psychological concept theory are prototype theory (cf. Rosch & Mervis 1975), exemplar theory (cf. Medin & Schaffer 1978), and theory theory (cf. Carey 1985). Though these theories have often also been discussed in the philosophical literature under the heading of the theory of concepts (see e.g. Fodor 1998; Margolis & Laurence 1999), Machery claims that “theories of concepts in psychology have in fact entirely different goals.” (op. cit.: 34) Effectively, Machery’s idea is that philosophers and psychologists work with job descriptions for concepts that are different enough to be independent, and so it should not be assumed that what psychologists have to say bears directly or even indirectly on what philosophers have to say when they use superficially similar terminology. Doing otherwise would be to commit an equivocation.

In response, critics (Edwards 2010; Margolis & Laurence 2010; Rey 2010; Weiskopf 2010) have argued variously that because concepts are constituents of thoughts (the dictum again) there is no need for Machery’s distinction, that the fact that psychologists

and philosophers intend to talk about concepts in a non-equivocal way shows the same, or that the distinction does not imply that the concepts psychologists and philosophers talk about are not the same. This last point is correct, but it is no grounds for insisting that the explanatory projects really *are* the same. The problem with the first and second responses is that they fail to be sufficiently sensitive to the point of the distinction that Machery draws. The point is that there is a question as to how these different explanatory projects are related. Neither the dictum nor the intentions of theorists is enough to show that they must be related in some particular way.

Though it is true that an approach to concepts that provides for a treatment of both of the two job descriptions distinguished by Machery would have the explanatory virtue of unification, Machery is surely right that it is not enough merely to *want* a unified explanation. Machery's point is best taken as the claim that there is a need to motivate the claim that these explanations fall together. To simply insist that they do is therefore dubious, both dialectically and theoretically. Referring to Prinz's (2002) list of desiderata, Machery makes the point that different criteria apply when evaluating the relative successes of theories addressed to them (2009: 36). One cannot simply assume that desiderata apply in the same way across different explanatory projects.

Having said this, it would be equally wrong, and a mistake Machery could perhaps be accused of making by omission if not commission, to proceed on the basis of the idea that the two explanatory projects should be treated separately. In fact, if we interpret the aim of philosophical concept theory as per the capacities answer, then any claim on this point must be withdrawn. The question must be left open. The explanation of referential coordination needs to come first, here as elsewhere; determining what is the relationship between psychological and philosophical concepts, and whether there is one, depends on what the best explanation of referential coordination demands.

Machery has relatively little to say about the relationship between the two forms of concept theory that he distinguishes, limiting himself to an argument against the idea that proponents of philosophical concept theory set tasks for psychologists to carry out (2009: 41-7), and some skeptical remarks about the role of thought experiments and intuitions about the attribution of propositional attitudes have in theorising about philosophical concepts. He sometimes gives the impression that he thinks the project of philosophical concept theory is suspect, because "we have no access to people's propositional attitudes except through the inferential ascription of propositional

attitudes,” (op. cit.: 48) and so it relies to a problematic degree on thought experiments and intuitive judgements. This is almost certainly incorrect. I can give you quite direct access to my propositional attitudes by telling you what I believe, what I intend, and my hopes and fears as well, and you can do this with your own thoughts by considering what you would say about the same. Carrying out explanatory projects concerning propositional attitudes, thoughts, does not need to rely on thought experiments or appeals to intuition. There is less murkiness here than Machery seems to suppose.

2.6 *Eliminating Concepts*

Machery uses the distinction between psychological and philosophical concepts to make room for an argument in favour of *eliminativism* specifically about psychological concepts. Eliminativism about psychological concepts is the idea that concepts are not a natural kind (2009: 230-246). This point is apt to be misunderstood. Psychological concept theory holds that there are things that have this feature. The point about natural kinds is not that this feature is not instantiated—that there are no bodies of information used as defaults in categorisation tasks—but that there is no interesting set of generalisations to be made about what instantiates it. Whether Machery is right about this (cf. Piccinini & Scott 2006; Weiskopf 2009) does not matter for present purposes. I mention Machery’s claim not because it will be a concern of this essay, but because there is a case to be made for a similar eliminativism, not contemplated by Machery, about *philosophical* concept theory, and this case turns on what we should conclude about referential coordination.

As noted above, philosophical concept theory is committed to there being entities that account for thoughts that are exercises of the same referential capacity also being exercises of the same conceptual capacity. This pattern of explanation can be made out in the various views of how to think about concepts discussed above: Peacocke’s stipulation that they are constituents of propositional contents and Fodor’s requirement that they be mental representations both take concepts to be either contents that determine referential content or vehicles that have referential content. Prinz includes the explanation of both intentional content (i.e. referential content) and cognitive content (i.e. patterns of sameness and difference in ways of thinking, roughly) as desiderata. The disagreement between those who identify concepts as contents and as vehicles takes place against an agreement on this point.

Philosophical concept theory is therefore committed to a *heavyweight* explanation of referential coordination in one form or another. But if the best explanation of referential coordination is a *lightweight* explanation, then concept theory is to that extent mistaken. It has a presupposition that should be rejected. There would be no kind of thing that would provide an answer to the identification question. This eliminativism about philosophical concepts is therefore congruent with Machery's eliminativism about psychological concepts.

The coordination functions explanation for which I will argue is lightweight. Since the argument will take the majority of the remainder of this essay to make out, it is too early to say that philosophical concept theory is in error. The point for now is just that it should be regarded as a live possibility that it is. That is, as long as we accept the idea that talk of concepts is best understood as offering an explanation of conceptual capacities, then it must be seen as making a substantive commitment that may turn out to be false. Machery remarks that “[o]nce at the center of philosophy, the philosophy of concepts has now been marginalized, maybe because for a few years now, it has been stalled.” (2009: 3) Machery's suggestion in the face of this is to “modify philosophers' relation to the psychology of concepts.” (ibid.) My suggestion is that the project might be better off being dismantled, and some parts thrown away.

In talking about concepts, I have been narrowly concerned with concepts of things, those that are involved in the exercise of referential capacities, rather than concepts of ways for things to be, such as concepts of properties or relations. The thesis of eliminativism about concepts therefore needs to come with a caveat. There is an obvious sense in which something similar to referential coordination goes for properties and relations; it can be rational to infer, from the thought that Cicero is a Roman and Caesar is a Roman, the conclusion that there are at least two things that are Roman. The lightweight explanation of referential coordination does not imply a lightweight explanation of this sort of coordination. I submit that it would be surprising if concept theory was wrong about things but right about ways for things to be; these would seem to stand or fall together.

It should also be emphasised that what gets called into question is concept *theory*, and *philosophical* concept theory in particular. Psychological theories of concepts are left untouched. It may also be the case that particular classes of entities identified as philosophical concepts still play a role in the best over-all theory of the mind; in

particular I have in mind the constituent mental representations that are vehicles of referential contents of thought. I will go into this further in §7.2.

2.7 Conclusion

Bringing in referential coordination and a handful of distinctions makes it possible to comprehend what talk of concepts is about. Concepts, in the philosophical literature at least, can be seen as theoretical entities that are linked to heavyweight explanations of referential coordination. Concept theory is therefore open to disconfirmation by the test of how well it explains referential coordination. The explanation of referential coordination does not, therefore, fall under the rubric of the theory of concepts, even though it is directly relevant to much of what does.

Chapter 3 - Sense Theory

3.1 Introduction

This chapter evaluates the sense-theoretic explanation of referential coordination. I will argue that, at best, sense theory can only provide an explanation that involves circularity; there is no way to establish that it is correct without already assuming that it is correct.

Sense theory finds its source in the semantical investigations of Frege, discussed in §1.3 above, as an attempt to build on his distinction between the sense and reference of an expression to provide a general account of rational content-involving phenomena such as thought and language. The sense theorist holds that thoughts are possessed of a sort of structured content distinct from and determinative of their referential content; these contents and their elements are senses. Sense theory is not the only way to develop Frege's notion of sense, and there are some internal disagreements which I shall discuss briefly. The basic idea behind the sense-theoretic explanation of referential coordination is that referentially coordinated thoughts are those that feature the same sense as an element of their content (§3.2).

In terms of the three questions, this puts us here:

	Psychologism	Referentialism	Weight
Senses	No	No	Heavy
Typed representations	Yes	Yes	Heavy
Mental files	Yes	Yes/No	Heavy
Coordination functions	Yes	Yes	Light

Because the sense-theoretic explanation locates the indication of sameness of reference in sameness of sense, it is adequate only if it can meet the *handles challenge*. This is the need to give some account of which thoughts have which senses as (part of) their contents. More exactly, the sense-theoretic explanation needs an account of what it is for a thinker to have the capacity to think thoughts featuring a particular sense, as distinct from any other. Such an account can be called a theory of *grasp*. If a theory of grasp can be given, there is still the possibility that it will make the overall sense-theoretic explanation fail to meet one of the explanatory desiderata (§3.3). To show that this second kind of objection

can be sustained, I will examine two representative theories of grasp: Chalmers' appeal to *epistemic intensions*, and Peacocke's appeal to *implicit conceptions* (§3.4). In both cases, the theory of grasp is adequate but only because it is circular. There is no way to motivate the theory of grasp, or say exactly how it is supposed to work, without presupposing the correctness of the theory in question, which means that the sense-theoretic explanation is also involved in circularity. This problem can be traced to the central commitments of sense theory, and the objection can be generalised. This motivates looking at psychologistic alternatives (§3.5).

3.2 *Developing Frege's Notion of Sense*

Frege's original introduction of the distinction between sense and reference inaugurates a number of complex issues about thought and language. Interpretative disagreements abound concerning how his notion of sense should be understood. Sense theory is one development of Frege's arguments and proposals. I will first bring out the core commitments of sense theory and how it provides an explanation of referential coordination. The notion of sense may be developed along either realistic or anti-realistic lines, or in constructive or non-constructive ways, though in considering how to apply Frege's notion of sense to the problem of referential coordination, we need only concentrate on realistic and constructive developments.

3.2.1 Core commitments and the sense-theoretic explanation

Chalmers (2002: 138-142) helpfully provides a brief statement of Frege's main claims about his notion of sense (the labelling is my own):

- (FS1) Every expression that has a reference has a sense.
- (FS2) Sense reflects cognitive significance.
- (FS3) The sense of a complex expression depends on the sense of its parts.
- (FS4) Sense determines extension.
- (FS5) In indirect contexts, expressions refer to their customary senses.
- (FS6) The sense of a sentence has an absolute truth value.
- (FS7) The sense of an expression can vary between occasions of use.

(FS1), (FS5), and (FS7) apply narrowly to language, and so can be put aside. (FS6) and (FS3) are both important, but need not concern us for present purposes. (FS2) and (FS4) will be the main focus, and provide the core of sense theory.

These two commitments can be expressed as the idea that there is a special kind of content other than referential content. Frege's preferred terminology was to talk about these contents as 'thoughts', though to avoid confusion with states and episodes of thinking, they can be called *Fregean propositions*. They are propositions because they determine a truth value, which Frege thought of as the reference of a thought. Fregean propositions are to be understood as complexes whose constituents are *senses*. Thoughts are themselves senses, but in what follows, 'sense' should be understood to mean constituent elements rather than whole contents.

Senses have identity conditions that reflect cognitive significance, as per (FS2). 'Cognitive significance', much like 'mode of presentation', is a bit of jargon that has accrued to the Fregean tradition, the idea behind which is not always clear. Roughly though, the idea is that sameness and difference of sense goes with what a thinker can rationally do with thoughts, such as employing them in inference or revising them in the face of new evidence. 'Goes with' is, of course, a phrase that offers little by way of illumination. The job (or one of the jobs) of the sense theorist is to give a fleshed out account of what this means. Fleshings out (FS2) will say something about how sameness and difference in sense determines or explains patterns of mental states and rational transitions between them. This elaboration will need to individuate senses in a way that satisfies (FS4). Because it depends on the individuation conditions on senses, determination of extension will not be accidental but rather *constitutive*: that there is a link between the two that explains cognitive significance is what makes a sense the sense that it is. We can call (FS2) and (FS4) the *individuation* and *determination* principles, respectively.

Contemporary formulations of sense theory are often put in terms of concepts. This sometimes takes the form of saying that concepts are individuated *at the level of sense* (Peacocke 1992). This is most naturally (though not necessarily) understood as the idea that concepts have the same individuation conditions as senses—which is just to say that concepts *are* senses. Of course, when taken this way, senses may not have or be able to do all that Frege originally wanted them to do.

This notion of sense links up in a fairly neat way with the explanation of referential coordination. Frege's original concern was to explain the closely related phenomena of what have come to be known as Frege's puzzle and Frege cases (see §1.3.1-2). The fact that referential coordination involves sameness of reference is naturally explained by the determination principle, and the idea of cognitive significance can easily be understood to include the rationality of the paradigmatic inferential dispositions. The way in which the individuation and determination principles can combine to provide such an explanation is reasonably clear. Thoughts that are referentially coordinated are those with propositional contents with senses as constituents. Sense determines reference, so if the propositional contents of two thoughts feature the same sense, then reference to the same thing is determined in the same way. If reference to the same thing is determined in the same way, then there is an indication of coreference.

As I will shortly argue, this explanation is in an important sense incomplete, though one can see the beginnings of a plausible explanation here.

3.2.2 Realist and anti-realist developments

Frege's notion of sense has been the topic of a large and contentious literature. Some of this can be understood as concerning the right way to develop it in order to do theoretical work. Others are better understood as concerning whether senses can be put to theoretical use at all. And sometimes these two concerns overlap. These disagreements often centre on the application of the notion of sense to questions in the philosophy of language, though they carry over to debates about thought as well.

One of the most important disagreements of the first kind is over the role of semantic properties, by which I mean the property of truth, reference (to things), and satisfaction (by things of predicates), which can be used to elaborate the contribution to truth of things at the sub-propositional or non-truth evaluable level. This disagreement applies directly to how the individuation and determination principles are to be elaborated. On one view, these properties are independent of the capacities that thinkers have to appreciate them. We can say that a thought is true, or that it refers to something, whether or not a thinker is capable of appreciating its truth, or knowing what it refers to, or what it is to satisfy it, in some non-trivial sense of those phrases. On the contrasting view, semantic properties have no explanatory role independent of the capacities for thinkers to make use of them (which is not to say that they have no explanatory role).

This disagreement is often characterised as a debate between *realists* and *anti-realists*. What is at issue beyond the rough characterisation just given is itself a matter of debate (cf. Brock & Mares 2014). In terms of sense theory, realists argue that semantic properties have an independent role in giving an elaboration of the individuation and determination claims, and anti-realists argue for the epistemically constrained notions in terms of warrant, justification, or evidence instead. Labels aside, the disagreement is usefully thought of as concerning the explanatory merits of the notion of sense that results. Anti-realists argue that appealing to properties like truth or reference is of no use in explaining the rationality of language or thought if it goes beyond any capacities that thinkers have to rationally respond to them (see Dummett 1959 for an early presentation of this view). Realists are moved to respond to this arguing that the epistemically constrained notions that the anti-realists take to be more explanatorily respectable cannot themselves be understood without appeal to semantic properties (cf. Peacocke 2008: 7-52).

This debate is a general one concerning the nature of content, though it may be difficult to state clearly what is at stake outside of particular applications. In any case, I propose to sidestep this debate altogether. The assumption I called Availability says that, in the core cases, referential content is unproblematically available. This assumption mandates looking only at realist versions of sense theory, for reasons of simplicity. Without the assumption, or on the anti-realist's view, it is unclear how the thoughts that need to be explained are to be identified, without doing so in terms of their referential content. This possibly accounts for (or perhaps it is the other way around) the fact that anti-realism often goes with a *language-first* approach to thought, the idea that one should come at the nature of thought via sentences in public language. This is insisted upon by Dummett (1991), for instance, and explicitly denied by Peacocke (1997). If one gets to thoughts via sentences in public language, then the problem goes away. But it is not clear that this can be done without presupposing a view of thought that implies a particular stance on referential coordination.

In any case, the assumption that referential content is unproblematically available in the core cases need not be unacceptable to anti-realists. They will simply accept it under sufferance, with the understanding that the core cases will need to be encompassed within a more general view acceptable on anti-realist lines. I will proceed as if things had been settled in favour of realism.

3.2.3 Constructive vs. non-constructive developments

A further disagreement amongst those who accept something like Frege's notion of sense concerns the extent and nature of the explanatory purposes to which the notion of sense can be put. Sense theorists both realist and anti-realist can agree that the point of appealing to Frege's notion of sense is to give explanations. Treatments of Frege's notion of sense that are supposed to be explanatory need to provide an account of what senses are so that the individuation and determination principles come out right, so as to be capable of providing *genuine* explanations; this needs to be done in terms that can be specified independently of sense theory. This means that they need to make use of the idea that, and so show how, they determine extension and reflect cognitive significance. This is what we could call the *constructive* character of sense theory. It aims to construct senses from materials independent of the notion of Fregean sense, so that they can be applied to explanatory problems that appear where things can be said to have content.

Non-constructive developments of Frege's notion of sense reject the general theoretical ambitions of sense theory. Such views do not necessarily posit entities in the manner of sense theory. For example, Sainsbury (2002) claims that while there are *relations* of sameness and difference in sense between token expressions, these relations do not need to consist of shared relations to entities that are elements of propositions. Sainsbury calls this a 'pared down' treatment of Fregean sense. It is explicitly limited to semantic theorising, and there is no question of entities that might be made use of in other projects.

Alternatively, sense might be held to involve more than just relations between words, but not things whose nature can be specified in independent terms. In this way, no non-circular account of what it is for something to have a given sense in its content can be given. The elaboration of how sense reflects cognitive significance will provide individuation conditions, and relations of sameness or difference sense, by *expressing* the relevant senses, rather than saying what the sense is, or what the relations of sense consist in. Non-constructive sense theory therefore offers no clue as to what senses are that does not presuppose the appropriateness of appeal to sense. McDowell's 'modesty' in the theory of meaning is of this sort (the term comes from Dummett 1975, himself a promoter of the opposed 'full-blooded' option). McDowell sees the point of sense *not* in providing theoretical explanations of content-involving phenomena, but in bringing

attention to a sort of datum that needs to be respected in the theory of meaning (something about which McDowell presents Frege himself as being somewhat confused, cf. 1998d: 190-1).

As with the debate between realists and anti-realists (with which this debate is closely tied), the question over the viability of a constructive treatment of Frege's notion of sense can be ignored for present purposes. Though they present not inconsiderable interpretive difficulties, it is possible to see the arguments McDowell brings to bear as lying very close to those that I will lodge against the sense-theoretic explanation of referential coordination (cf. the argument against Peacocke in McDowell 1994: 162-174, and the dilemma for full-blooded theories of meaning presented in 1998b, 1998c). The argument of this chapter will effectively be that the case against a constructive treatment of Fregean sense is correct. But whereas McDowell seems to want to derive from this a sort of anti-explanationist moral (cf. the argument against 'sideways-on' views of concepts, to the effect that no non-circular theory of the contents of thoughts can be given in McDowell 1994), I take it that this a reason to look at alternatives to sense theory.

3.3 *The Handles Challenge*

The sense-theoretic explanation of referential coordination makes use of the idea that thinkers with referentially coordinated thoughts have capacities of a particular sort. These are capacities to think thoughts with particular senses as constituents of propositional contents. An initial worry that one might have about this as an approach to the problem of referential coordination is that it involves a very substantial set of commitments. Concentrating on that problem only, as I have argued that we should, it is not obvious why such commitments should be necessary. While some may wish to respond by arguing that treatments of problems, such as referential coordination, that turn on rational content-involving phenomena need to make use of a more fine-grained notion of content, this response is not particularly compelling. Why must content be the right place to locate the explanation? A different response might be to say that, even if we take referential coordination as an independent problem, it promises a theoretical approach to that problem that would unify it with solutions to a range of other issues to do with rational content-involving phenomena, and that is something desirable even if not mandatory. Given that this is an attraction that sense theory has had for those who advocate it, the initial worry needs to be set aside.

If this response is somewhat more compelling, then it simply raises the question as to whether the sense-theoretic explanation of referential coordination is a good one. Whether it is a good one or not depends on how the explanation sketched above gets completed. That sketch made use of relations of sameness and difference between thoughts in terms of the senses that constitute their content. So there needs to be some basis for determining which senses appear in the contents of which thoughts, and so for determining which capacities a thinker has that can figure in the explanation in a given case.

We can call a thinker's possession of a capacity to think thoughts that feature a particular sense *grasp* of that sense. An account of what makes for such capacities we can call a *theory of grasp*. Though this terminology may be familiar, the use of 'grasp' here is slightly unhappy, as it is also the term used for how language users understand expressions (one grasps their meaning) and so it can obscure the difference between the semantic and the psychological explanations that sense theory offers. I will stipulate that grasp of a sense is to be understood only and strictly to mean possession of the capacity to think thoughts featuring that sense.

With this terminology, we can state the following challenge:

Handles challenge: A sense-theoretic explanation of referential coordination is incomplete without, and so requires, a theory of grasp.

It has sometimes been suggested that something like this suffices as an objection to sense theory without needing to look at what theory of grasp might be given. So it has been claimed that, since senses are mind-independent abstracta, what it would be to be related to a sense in the 'grasping' sort of way is *obscure*, or else that there might be *multiple* ways of being related to a sense in the 'grasping' sort of way and the only way to avoid this is by making an arbitrary stipulation (Fodor 1998: 17-20; Margolis & Laurence 2007: 580-1). Neither sort of argument is compelling. It is not at all obvious why senses being mind-independent or abstract makes it obscure what it is to grasp them. By the sense theorist's lights, senses do play a role in psychology, so simply insisting that they cannot sounds like uncharitable interpretation at best. And their playing some role in psychology is consistent with their not being themselves reducible to psychological entities. Nor is it obvious why being abstract or mind-independent should mean that the sense theorist can

only rule out multiple ways of being related to a sense in the ‘grasping’ sort of way by means of arbitrary stipulation.

The sense theorist can make a more direct response by offering some elaboration of what being related to a sense in the ‘grasping’ sort of way is. Here is Burge:

Grasping a thought [i.e. a Fregean proposition] is simply a misleading metaphor. Any view should cash out the metaphor in terms of having a certain ability to think. Such an ability is attributable on the basis of ordinary evidence and is *constitutively associated with a variety of applicational and inferential abilities*. How does one “grasp” a thought content? One thinks it. Even on a Platonist view [i.e. one that takes senses to be mind-independent in a particularly strong sense], the thought contents should be regarded as playing a role in type-individuating mental events, states, and abilities. On any view, the contents should be regarded as abstract, in order to account for the multiplicity of instances of events that they type-individuate, to account for the shareability of kinds of thoughts, to account for compositionality, to account for the structure of inference and for various aspects of truth and justification. (2004: 30; emphasis added)

Burge is talking about senses at the level of whole propositions, but the point goes for constituent senses just as well. The idea is that what it is to grasp a sense is just whatever it takes to have an ability to think thoughts that have propositional contents that feature that sense. This suggests that the fact that senses are abstracta, whether mind-independent or otherwise, does not itself pose any special problem for their playing a role in psychological theorising. If their role was meant to be causal, then there would be a problem. Burge’s suggestion that their role is that of type-identification, or type-individuation, avoids this, since that role is not causal.

Allowing that what it is to grasp a sense is to be able to think thoughts of a particular kind does not mean the handles challenge is resolved, however. Easing worries arising from the ontological status of senses does not remove the need to say what is involved in grasping a sense. The point of the handles challenge is that there is a need to say *why* it is a good idea to attribute to thinkers capacities that are to be type-identified in terms of

senses. Without this, there is no explanation. The worry about arbitrariness remains live. This presents the potential for two sorts of objection.

A theory of grasp is *adequate* only if it shows that there is not more than one way to grasp a sense (at least, at a level relevant to the kinds of explanations that senses might provide; differences at the of level chemistry, for instance, would not provide for different ways of grasping senses). An inadequate theory of grasp will not demonstrate this. The potential objection is that there is no adequate theory of grasp, and so the appeal to sense must be rejected as a candidate explanation of referential coordination. Call this the *adequacy* objection.

Supposing a theory of grasp is adequate, it may involve commitments that have the consequence that the sense-theoretic explanation fails to meet the desiderata on an explanation of referential coordination. There might be some appeal to facts or conditions that apply only in limited cases; the theory might attribute psychological capacities to thinkers there is no independent reason to think that they have; or the theory might illicitly make use of resources that are only available on the prior assumption that the theory is correct. These all present ways sense theory may fail to meet one or other of the desiderata on a theory of referential coordination. Call this the *explanatory* objection.

Constructive realistic sense theory would appear to be well placed to respond to the adequacy objection. They aim to provide an account of what senses are which can make use of the referential properties of thoughts, and an account of what makes for sameness and difference in sense can be provided on that basis. A theory of grasp can be given by specifying the properties of thinkers that relate them in the 'grasping' sort of way to a sense individuated according to the constructive account. This would show that, for each distinct sense, there is one way of determining referential content, and so one way of grasping it, i.e. one way of having the ability to think thoughts whose referential content is determined in that way. Such an account would provide a demonstrably adequate theory of grasp. However, it is not a given that adequate theories avoid the explanatory deficits just mentioned. Ultimately, critics of sense theory who think that senses cannot provide an explanation of psychological phenomena are correct, though strictly speaking, the conclusion must be that they *can* explain psychological phenomena, but it will not be a *good* explanation. But in order to see this, it is necessary to look at some theories of grasp that have actually been given.

3.4 *Theories of Grasp*

This section sets out two approaches to sense theory and why they provide an adequate theory of grasp. I provide (but do not endorse) some motivations for the theory in each case.

3.4.1 Epistemic intensions

One way to approach giving a theory of grasp is to do so in terms of the ‘variety of applicational and inferential abilities’ that Burge mentions. To say what ability can be type-identified by means of a determinate sense, on this approach, is to pick out a specific and distinctive set of possible manifestations; call this a *dispositional* approach to the theory of grasp. The dispositional approach can be motivated by a particular sort of proposal about how to think of senses. The association with the applicational and inferential abilities can be thought of as representational, in the way that formal objects, such as mathematical entities like functions or vectors, can be used to represent non-formal phenomena like traffic flows and magnitudes of physical force. This representational association is direct, and so is the relationship between a sense and what it is to grasp it. There can be no complete overlap in the dispositions of thinkers relevant to senses that are nevertheless distinct, since there is nothing more to a sense than the representation of those dispositions. Sameness and difference in sense is just sameness and difference in the relevant dispositions.

A detailed treatment of a dispositional approach to sense theory can be found in the work of Chalmers (2002, 2006, 2011, 2012), and I will take this as representative.

At the heart of Chalmers’ complex treatment is what he calls *epistemic intensions*. These are an extension of the intensions familiar from Carnap (1947). A Carnapian intension, or *subjunctive* intension, is a sort of mathematical object: a function from possible worlds to extensions. Carnap’s purpose was to use this to do work in semantics and modal logic. In the case of sentences, the corresponding subjunctive intension is a function from possible worlds to truth-values; in the case of words, the corresponding subjunctive intension is a function from possible worlds to individuals, properties, relations, classes, and so on. Sentence-level intensions can be thought of as structured complexes with word-level intensions as parts. Subjunctive intensions bear similarities to Frege’s notion of

sense, since they determine extension and can form structured complexes. But they lack the connection to cognitive significance distinctive of senses.

Epistemic intensions reinstate this constitutive connection. Instead of functions from possible worlds to extensions, they are functions from *descriptions* of how the world might be, which Chalmers calls *epistemically possible scenarios*, or just *scenarios*, to extensions. An epistemic intension, in combination with how the world in fact is (i.e. which scenario is actual), determines a subjunctive intension. Chalmers (2006) puts this in terms of *primary* and *secondary* intensions. Primary (epistemic) intensions account for facts about rationality, while secondary (subjunctive) intension account for facts about modality. This *two-dimensional* aspect (elaborated at length in Chalmers op. cit.), though crucial for the applications Chalmers makes of the theory, can be set aside.

In what way do epistemic intensions account for facts about rationality? On this way of constructing senses, a sense is considered to be an epistemic intension, and distinct senses are distinct epistemic intensions, on the understanding that the individuation of senses goes with characteristic applicational and inferential abilities that epistemic intensions represent. The specific inferential abilities are those manifested in a set of dispositions to infer from a scenario to an extension (a truth value in the case of structured epistemic intensions, an individual, property, relation, or class in the case of elemental epistemic intensions). A thinker can think a thought with a certain cognitive significance—a particular sense in its content—if she can think thoughts that she would be *a priori* disposed to infer from a particular set of scenarios. Crucially, this disposition is relative to certain idealisations. A thinker's actual inferential dispositions may go wrong in various ways, and the relevant set of dispositions are those where these mistakes are absent.

This can be summarised like so:

Epistemic intensions theory: A thinker grasps a sense if and only if they have idealised *a priori* inferential dispositions that are represented by the relevant epistemic intension.

Epistemic intensions theory provides a demonstrably adequate theory of grasp. Epistemic intensions theory implies that there is a one-to-one correspondence between senses and epistemic intensions, since there can be only one condition, one set of idealised inferential dispositions, sufficient for grasping a given sense. There is no set of idealised

inferential dispositions sufficient for grasping a sense that cannot be represented by the relevant epistemic intension, and so there cannot be multiple ways to grasp a sense that involve different patterns of cognitive significance.

Though epistemic intensions theory is not merely stipulative, the explanatory objection to theories of grasp might nevertheless be pushed, somewhat in the spirit of Fodor's objection, on the grounds that there are no independent reasons for thinking that what the theory of grasp presupposes as a way of constructing senses really obtains, beyond the way in which it can feed into a theory of grasp. In the present case, the worry would be that there is no motivation for idealised inferential dispositions that can be represented as epistemic intensions. Not just any set of properties for which there is a one-to-one correspondence with a way of constructing senses will do by way of an adequate theory. So what independent reason is there for accepting the epistemic intensions theory?

One consideration that can be offered is this:

Constructive Basis: We form judgements regarding extensions about cases. Forming a judgement about a case is an inference from a description. If one idealises the descriptions to get scenarios, and idealise the inferences that are made so that they are purely rational, then one gets epistemic intensions.

The capacity to make extension-identifying judgements automatically grounds epistemic intensions by providing the materials for their construction. Another set of considerations by way of motivation concerns the advantages of the epistemic intensions theory:

Construction Without Speculation: Constructing senses by way of epistemic intensions provides a theory of grasp that avoids tendentious speculations about the psychological states of thinkers. In particular, it provides a theory of grasp without imputing inner states that may or may not be realised. By contrast, the dispositions that can be represented by an epistemic intension can be multiply realisable, so the theory of grasp that results is not committed to any speculations about psychology.

The elements involved in the construction are simply those required by the data that give rise to Frege puzzles in their various guises, and referential coordination more generally, both of which concern the rational properties of thoughts that can be represented by epistemic intensions. Nothing has been brought in that we did not have reason to think was present in the set up of the case.

3.4.2 Implicit conceptions

An alternative to the dispositional approach instead looks to the *basis* of inferential dispositions. The relevant inferential dispositions in either case are those that serve to identify an extension. So on this alternative, senses are to be constructed from conditions for determining reference; different conditions for determining reference correspond to different senses. This sort of approach is represented by the treatment provided by Peacocke (2008, developing his 1998). In Peacocke's terms, the conditions for determining reference that individuate a sense are what he calls *fundamental reference rules* (FRRs). These are fundamental, because while there might be multiple ways for picking out the reference of a given expression or thought, this will be true in virtue of them and not the other way round.

Peacocke places great emphasis on the fact that the dispositions in question are rational ones. He comments (2008: 114n2) that his earlier and perhaps better known (1992) theory of grasp, which was given in terms of 'primitively compelling inferential dispositions' and hence is much more in line with the dispositional approach, is inadequate to account for this. On this point, given that we are interested in a sense-theoretic explanation of the rationality of a certain form of inferential disposition, those important for understanding referential coordination, it seems appropriate to concentrate on not just any basis for the dispositions, but on what we can call the *rational* basis; call this the *rational basis* approach.

Although the central claim of the rational basis approach offered by Peacocke is the idea that senses are individuated by FRRs, this is meant to be closely tied to considerations to do with what it takes to grasp a sense. The details are developed by looking at some specific examples (2008: 56-76), of which four are particularly important (I follow Peacocke's preferred terminology of 'concepts' rather than 'senses'): perceptual or observational concepts, logical concepts, demonstrative concepts, and the first-personal concept. As an example, for "the observational concept *oval*," we get the following:

What makes something fall within the extension picked out by the observational concept *oval* is that it is the same shape as things are represented to be in perceptual experiences of things as oval. (This representation will be of the correctness of a non-conceptual content in my view). (op. cit.: 56-7)

This FRR features what Peacocke calls an ‘objective-perceptual’ component, that identifies a feature of the world, and an ‘identity’ component, that makes that feature apply to things more generally (op. cit.: 31).

The connection between an FRR and what it is to grasp a sense is that senses individuated by FRRs are components of propositional contents of thoughts that are subject to certain norms, and these norms are meant to be explained by the FRR. This means that, if one thinks a thought in a way that involves one’s grasp of a particular concept, then one will be subject to a particular set of requirements in one’s thinking that thought. These have bearing on their ‘introduction’ and ‘elimination’: only certain other mental states will rationally entitle one to form that thought, and having formed it, one will thereby be rationally entitled to make transitions to certain other thoughts. These norms articulate the analogue of the ‘inferential and applicational abilities’ that type-identify thoughts featuring a determinate sense. Senses individuated by FRRs are therefore constructed from *requirements on* rational inferential dispositions, rather than those dispositions themselves.

Peacocke substantiates the idea of there being such requirements by employing what he calls *argument types* (op. cit.: 61-72). These show that one can derive FRRs from norms, and norms from FRRs, in at least a few and possibly many different ways. In order for FRRs to figure as the rational basis for inferential dispositions, Peacocke claims that a kind of mental state that has the FRR as its content must be present in the thinker’s psychology. These are what he terms *implicit conceptions* (op. cit.: 113) They are implicit, because it is often not the case that a thinker can articulate their content in a way that would constitute their being explicit. They are conceptions because they involve an identification of a referent in a certain way, one that corresponds to the conditions given by a FRR.

The theory of grasp we get from this can therefore be stated like so:

Implicit conceptions theory: A thinker grasps a sense if and only if they have an implicit conception of the fundamental reference rule that individuates it.

Implicit conceptions theory can provide a demonstrably adequate theory of grasp, since for any given sense, there is only one way to grasp it, by having the relevant implicit conception.

As with epistemic intensions theory, implicit conceptions theory is in need of independent motivation. Neither of the two points motivating epistemic intensions theory go for implicit conceptions theory, as one would expect given their quite different commitments. The sorts of considerations Peacocke adduces have to do with the norms that he argues apply to thoughts that feature particular senses (op. cit.: 72-6):

Distinctive Norm Explanation: The norms that apply to thoughts that feature a given sense can be explained by the fundamental reference rules for that sense.

A closely related proposal (op. cit.: 134-40) is what Peacocke calls

Phenomenon of New Principles: The rationality of a thinker's acceptance of principles governing a sense that do not follow from any principles the thinker already accepts can be explained by the fundamental reference rules for that sense.

Concepts—senses—are constitutively and definitionally tied to rationality in this way. The thesis that a concept so understood is individuated by its FRR is a substantive, non-stipulative philosophical thesis about such concepts. On this conception, only certain rules will be capable of explaining norms, and the acceptance of principles, that are distinctive of concepts. It is these reference rules which, as a matter of substantive fact, individuate certain senses (op. cit: 60).

3.5 *Circularity Problems*

There are clearly responses to the handles challenge that at least avoid the adequacy objection. But sense theory is not in the clear. The explanatory objection is still there. In this section, I will argue that there are explanatory problems with both theories of grasp: they make the sense-theoretic explanation circular in two ways, circularity in how the theory is motivated, and circularity in how the theory is applied. Different though they are, the problems are parallel for both theories. Motivating and applying the theory of grasp requires assuming the correctness of the theory; not doing so simply brings us back to the handles challenge. The fact that these are parallel problems indicates a common root in general commitments of the sense-theoretic explanation, and I will use this to generalise the objection beyond the particular theories considered.

3.5.1 Circularity problems for epistemic intensions theory

Epistemic intensions theory relies on the idealisation of inferential dispositions. As such, it needs to give some elaboration on *which* idealisations are the relevant ones, and it is not clear that it can do this without making the account circular. How can the *non-ideal* dispositions be identified? And, how should the non-ideal dispositions be *made* ideal? Answers need to be given that do not simply assume the correctness of the theory. And such answers are not available.

The first question points to a problem of circularity in how the theory is motivated:

CM1: There are no grounds for determining which inferential dispositions are ideal and which are non-ideal that are independent of which epistemic intensions represent them.

What it means to have *non-ideal* inferential dispositions of the relevant sort is that the thinker's actual dispositions will sometimes pick out the wrong extension, or the right extension in the wrong way, by involving some sort of mistake or mis-step in her reasoning from scenarios to extensions. What determines which of a thinker's inferential dispositions are non-ideal? The answer must be that inferential dispositions are non-ideal if they deviate from the inferential dispositions that would be represented by the epistemic intension of the relevant sense were they idealised. But this makes the account circular, since we needed to know which is the relevant sense. And it is not clear what alternative there might be.

This point undermines a presupposition of what I called Constructive Basis, the idea that representing idealised inferential dispositions with epistemic intensions is simply a matter of doing some work on inferential dispositions that thinkers can be said to have anyway prior to the theory. In fact, things are the other way around: the inferential dispositions have to be understood as those that can be idealised by means of epistemic intensions.

The epistemic intensions theorist could respond by arguing that the class of non-ideal inferential dispositions is defined by the presence of ordinary faults in reasoning. Ordinary faults include occasions when one forgets a step, or makes a mistake about what follows from what, or about the distribution of one's prior probabilities. But what is being idealised are inferential dispositions, and presumably mistakes in reasoning are errors in the *manifestation* of an inferential disposition, not in the dispositions themselves. So the absence of ordinary faults in reasoning does not define the class of non-ideal inferential dispositions.

Even if this were sufficient, appending this to the theory is unsatisfactory without an answer to the question of *why* the class of non-ideal inferential dispositions is defined by the presence of ordinary faults in reasoning. This cannot be something that is available outside of the theory, but must have something to do with the fact that ordinary faults make inferential dispositions not candidates for representation by an epistemic intension. In addition, there is a question about whether ordinary faults in reasoning really are all that goes into defining the class of non-ideal inferential dispositions. The theory cannot exclude there being extra-ordinary faults in reasoning that could produce deviation in the extension. The only obvious way of ruling this out is to say that ideal inferential dispositions are those that can be represented by the relevant epistemic intension, so that ordinary faults in reasoning define a class of non-ideal inferential dispositions because they cannot be so represented. With this, we are back with the question about how to understand the ideal/non-ideal contrast. It appears to depend on what it is supposed to explain.

The second question points to a problem of circularity in application:

CAI: For any given non-ideal inferential disposition, there are no grounds that determine how to idealise that are independent of which epistemic intension represents it.

This problem also turns on the reliance on idealisation. How should a thinker's non-ideal inferential dispositions be idealised? Any non-ideal inferential dispositions can be idealised in more than one way, since there will be more than one way to correct the process of reasoning so as to determine an extension. Any putatively non-ideal inference therefore presents a choice between maintaining the extension determined by the faulty inferences, or not, and between different ways of modifying the extension. Which idealisation is the right one to make must be answerable to which sense she is supposed to grasp, since what inferences one is disposed to make is meant to be captured by the senses that figure in the thoughts that one has the capacity to have.

Differences in how inferential dispositions are idealised will make a difference to which sense is grasped, because they will make for difference in which epistemic intension represents them. In which case, this theory provides no way of determining which idealisation should be made without presupposing that a particular sense is being grasped (that is, that the thought is represented by a particular epistemic intension). And if it cannot be determined which inferences are the ideal ones without presupposing grasp of a particular sense, then epistemic intensions theory would have the consequence that one cannot say, for any given sense, whether or not a thinker grasps that it without already presupposing an answer. Hence applying the epistemic intensions theory of grasp involves the sense-theoretic explanation in a circle.

The inferential dispositions are supposed to be *a priori*. Perhaps it could be objected that this provides a way out of the circle. Some sort of rationalist epistemology is a natural counterpart to the Fregean view of content, and the support it might lend to or draw from such an epistemology is often highlighted by advocates of the Fregean view of mental content, not least Chalmers himself (cf. his 2004; 2006; see Schroeter 2004 for critical discussion). Idealisation, it might be said, is only an extension of what happens when, attending carefully to how we are reasoning, we consider what we would count as being the same individual, or instantiating the same property. This is one way of construing Constructive Basis. Putting aside the question of whether that idea is sound, the epistemic intensions theorist might argue that certain inferential steps will strike us as obviously right or as obviously wrong, where that is not informed by any information we have about how things are with the world that we could find out by going and looking. Circularity in application is avoided because *a priori* reflection is something we do have reason to think is available to us independent of the theory. Although we might not be

able to say independently of enjoying this *a priori* reflection *which* idealisations apply, we can say that they do apply.

An initial worry about this response is that the idea of *a priori* reflection goes far beyond what could reasonably be supposed falls under the motivation given as Constructive Basis. It makes the account psychologically unrealistic, since there can be little reason to credit thinkers with an epistemic power directed at their own inferences that has no parallel directed at other domains. Aside from this, as a response to the circularity problem, the appeal to *a priori* reflection is undermined by the fact that thinking that the correct way of idealising inferential dispositions is something to which one can have access through *a priori* reflection, is not independent of accepting the theory. *A priori* reflection would provide access to how to idealise an inferential disposition only if it is already supposed that such reflection is based on the senses that are grasped. By itself it provides no indication that thoughts do involve grasping senses, much less that they are representable by epistemic intensions. The circularity problem arises again when cashing out what *a priori* reflection would need to be.

I conclude that a theory of grasp on the lines of epistemic intensions theory would involve the sense-theoretic explanation in vicious circularity.

3.5.2 Circularity problems for implicit conceptions theory

Implicit conceptions theory will not have exactly the same problems for exactly the same reasons, because idealisation does not feature. What gets posited in implicit conceptions theory is a state that is present whether or not it is manifested in one way or another in a set of dispositions. However, implicit conceptions theory has similar problems for similar reasons. It makes use of the idea that an implicit conception *explains* a set of inferential dispositions. Just as epistemic intensions theory needs to constrain the idealisations of the relevant inferential dispositions, implicit conceptions theory needs to give content to the idea that implicit conceptions offer this sort of explanation, and this need gives rise to circularity in a parallel way.

Implicit conceptions theory is motivated by the thought that there are norms governing the use of certain concepts, and that these norms need a particular sort of explanation. But this is not really an independent motivation for the theory. One needs to already assume the correctness of the theory in order to generate the data to be explained:

CM2: There is no reason to accept the existence of the norms governing thoughts featuring particular senses, or principles that are rationally accepted, unless one already accepts the individuation of senses by FRRs.

If this is right, then the fact that the alleged norms in need of explanation are those *distinctive* of thoughts featuring a particular sense means that there is no way to motivate their existence independent of the idea that there are senses individuated by FRRs. So the explanatory considerations provide no motivation without already assuming the correctness of the theory being motivated.

The defender of implicit conceptions theory might want to appeal to *a priori* applicability of norms as a way of giving the explanation some content, in the rationalist style considered in the context of epistemic intensions theory. The idea would be that the norms can be formulated independently of any theory because they are obvious or at least available to rational reflection to any thinker sufficiently equipped with the cognitive wherewithal to do so. But, as before, it is unclear that this is a legitimate fix. The norms that Peacocke thinks apply *a priori* only apply at all if the concepts in question exist. Otherwise, they may be different norms, which may have different implications.

At this stage, it is worth raising a general worry about implicit conceptions theory. The idea that FRRs are brought out by reflection on the norms governing thoughts can be undermined by considering alternative *deflationary* explanations that do not involve implicit conceptions. Such deflationary explanations will locate what implicit conceptions theory holds to be the content of an implicit conception—call this a *putative FRR*—in something other than a thinker's mental states. On a deflationary explanation, rather than being *implicit in* a thinker's mental states, putative FRRs are implied by something else, and what that is will involve things other than the sense in question in systematic ways.

Deflationary explanations apply to many of the examples that Peacocke provides. As he acknowledges (2008: 122), in giving the motivation for implicit conceptions, Peacocke often relies on examples where the referents are mathematical and logical entities: arithmetical functions, such as the concept of a limit of a sequence, or truth functions, such as the concept of disjunction. Such concepts refer to *formal* entities. Formal entities

occupy a place within a formal system, such as ways of filling out a truth table, or operations on topological spaces. Such entities admit of definitions. These are canonical ways of spelling out their place within such systems. They compare, not incidentally, with frequently given examples of definitions involving gender and kinship relations, like bachelorhood (=def unmarried man). But given their occupation of places within a formal system, giving the conditions that define this sort of entity can be understood as a matter of finding out something implied by the place of the referent in the relevant formal system, as opposed to finding out something implicit in the mental states of thinkers. On the deflationary view, putative FRRs for mathematical and logical entities are not reference rules but claims about formal entities and their place in formal systems. They reveal something about these special sorts of entity, not something interesting about the psychology of thinkers that can think about them.

In the other sorts of examples, the putative FRRs seem to be statements of generalisations about the use of expressions in public language. These generalisations have little to do with the machinery posited by implicit connections theory. In the case of indexical and demonstrative concepts, putative FRRs are generated by providing an explicit statement of how such words are understood to refer to elements that are salient or have positions in a context, for example, the speaker or the time of a context. In the case of observational concepts, the FRRs are generated by introducing a bit of neologism stipulated to refer to a theoretically posited 'non-conceptual content' (cf. op. cit.: 56-7). The deflationary explanation in these cases is that putative FRRs are implied by these (stipulative) semantic facts, rather than what is implicit in the mental states of thinkers.

There are, admittedly, examples Peacocke discusses that do not fit this pattern. But these consist of what must be considered inadequate attempts to provide FRRs. One such example is given as an example of explanation by implicit conception, concerning what he calls 'the concept *chair*':

One of the thinker's perceptual systems, say, identifies some object in the environment as having a supporting area and a back, and the subject has the background information that the object is used for sitting on. This information from the perceptual system, together with the background information, is combined, at a subpersonal level, with the content of the implicit conception involved in possession of the concept *chair*. It is computed, from this body of information, that the presented object is a

chair. This in turn explains the thinker's willingness to judge that that object, demonstratively given in perception, is a chair. (2008: 140)

The putative FRR for *chair* seems to be “having a supporting area and a back, ... used for sitting on.” This is clearly not sufficient for picking out chairs as distinct from other sorts of thing. Sofas are not chairs, but they have backs and supporting areas, and are used for sitting on. Interestingly-shaped boulders have supporting areas and backs of a kind; does someone’s using them for sitting convert them into chairs? Perhaps chairs are artefacts rather than natural objects—perhaps you and I have different opinions on that. Does this show that we understand the concept *chair* differently while still thinking about chairs as such? Or does it show that at least one of us has failed to grasp it ‘properly’? There is really no way of saying, at least not without already buying into the theory that is under scrutiny. The implicit conceptions theorist is committed to the idea that one of us must fail to grasp the concept, or fail to do so properly. But whether it makes sense to appeal to the idea that people can fail to properly grasp a concept is what is at issue, including as it does the idea that there is anything to grasp at all. So that can add no weight to the proposal.

More importantly, for our purposes at least, it is very hard to see what putative FRRs for individuals could be when they do not involve specifying a relation in which the thinker stands to that individual, as in the demonstrative or indexical case. It is hard to know what an implicit conception would be other than a description, and these are notoriously unreliable for picking out correct extensions (cf. Kripke 1980). Even though implicit conceptions need not have the form of descriptions, it is not encouraging to see Peacocke treating the Hesperus/Phosphorous case with the idea that the implicit conception of Hesperus is that it is the brightest star in the evening sky (2008: 60).

Peacocke remarks that

[s]ince it can be hard to make explicit the content of one of one's own implicit conceptions, we should equally not be surprised if thinkers sometimes mischaracterize the content of their implicit conceptions. A thinker's explicit endorsement of an incorrect definition does not mean that he does not have an implicit conception whose content is the correct definition. (2008: 122)

This seems to be correct as far as it goes. But the point about cases where the putative FRR is inadequate, such as that given for *chair* or for concepts of individuals thought of in non-indexical or non-demonstrative ways, is that the deflationary explanation applies here as well, albeit in a different manner. In cases where the putative FRR is *not* inadequate, there is an alternative to the explanation to do with either facts about the referent, as in the mathematical and logical case, or facts about semantic properties of certain special expressions in natural language. The deflationary explanation of cases where there is difficulty in stating an adequate FRR, or where it is only possible to offer a vague statement, is that it is difficult because the sort of basis for putative FRRs that the deflationary explanation exploits is missing; there is no formal system in which something is located, and no special linguistic expressions about which to generalise.

Might it be that, like epistemic intensions, FRRs for individuals are description-*like*, though highly complex? This would predict the difficulty in stating correct FRRs for individuals. But it does not help matters to say that implicit conceptions are *like* descriptions, in that they determine an extension, but different, without then specifying what that difference comes to. It is also unclear that highly complex FRRs for individuals can reasonably be attributed to thinkers capable of the simple task of thinking about individuals. This seems to stretch psychological realism to the limit.

Peacocke says that “[t]o make an implicit conception explicit can, then, on occasion be a major intellectual achievement.” (op. cit.: 120) For the examples he gives, what effort is needed seems to be directed more towards finding out facts about how things are, either with the logical or mathematical aspects of the world, or with semantical generalisations about words. It is not clear that this deflationary explanation lacks something that implicit conceptions theory provides. The problem of circularity in motivation aside, this point undermines the introduction of the class of mental states that implicit conceptions theory needs.

The role of explanation in implicit conceptions theory is also at the root of the problem of circularity in application. Despite it being an important part of his overall view of thought, Peacocke actually says relatively little on the details of what exactly the explanatory link between the rationality of thought and inference and implicit conceptions is supposed to be, and at one point issues a promissory note for a fuller account in later work (op. cit.: 146). As Peacocke says, however, “explanation by an implicit conception is a species of explanation by a content-involving state, the content

being the content of the implicit conception.” (op. cit.: 140) However the exact details are to be worked out, the content-involving nature of the explanation will need to be preserved. This brings up the following problem:

CA2: If explanation by implicit conceptions is a species of explanation by content involving state, then implicit conceptions figure as premises in inferences; implicit conceptions can figure as premises in such inferences only if the content of the implicit conception features the sense in question; in which case, the explanation is circular.

The problem can be illustrated by one of the examples Peacocke supplies for proposed implicit conceptions. For the sense he calls ‘the observation concept *oval*’, the implicit conception is this:

What makes something fall within the extension picked out by the observational concept *oval* is that it is something of the same shape as things are represented to be in the perceptual experiences of things as oval.

If grasp of the sense picked out by *oval* consists in knowing this proposition, then the account of what it is to grasp this sense presupposes the thinker’s grasp of it. Suppose *oval* did not appear in the implicit conception; there would then need to be some explanation of why this implicit conception explains the rationality thoughts and inferences it is supposed to. The account cannot just rely on the assertion that the implicit conceptions do have this explanatory role. It needs to say something about how they do. And the problem is that a necessary condition on doing this is that the implicit conception features the sense the grasp of which which it is meant to provide.

Admittedly, not all of the FRRs Peacocke supplies are like the that given for *oval*. But this only means that the FRRs he supplies will not all be adequate to providing an explanation. If the sense does not figure in the content of the implicit conception for it, then there is no way, certainly no obvious way, for the implicit conception to be explanatorily relevant to the relevant thoughts and inferences. So the account seems to be adequate only on the condition of being circular.

A response might be to argue that the implicit conception does not need to feature the sense in its content, but rather only needs to refer to the sense. More exactly, the sense

figures in the implicit conception as what the implicit conception is *about* rather than in its content. Implicit conceptions are, it could be said, conceptions about particular senses, rather than the referents of those senses. This is enough to make them explanatorily relevant without being circular.

However, implicit conceptions referring to senses in this way would have the consequence that implicit conceptions theory is committed to more than implicit knowledge of FRRs, but also implicit knowledge of how rational thought works. It would hold that thinkers implicitly know that there are such things as senses, and that they are individuated by FRRs. In other words, rational thinkers would implicitly believe in Peacocke's theory (coincidentally guaranteeing that the theory is correct). Aside from being scarcely credible in itself, this move would have the consequence that implicit conceptions theory is saved from circularity only at the cost of attributing capacities that there is no reason, independent of the theory, to attribute to thinkers capable of referentially coordinated thought. This counts as no less an explanatory deficit, and so does not do much to preserve what might be otherwise appealing about the theory.

The problem stated in CA2 is, in a way, recognised by Peacocke (op. cit.: 144-8), though he does not go so far as to say that circularity will be part of *any* implicit conception adequate for giving explanations of rational inference. The way he puts it is by acknowledging that implicit conceptions theory involves a departure from the account given in his earlier *A Study of Concepts* (1992) because it drops the requirement of what he there called the 'A(C) form of possession conditions for a concept' (ibid.: 9); that is, a requirement on the form a theory of grasp for any given sense will take. The A(C) form requires that the sense not figure within the scope of a propositional attitude; that is, in the content of thoughts, of the thinker who grasps it as a condition on their doing so. Dropping this requirement allows circularity to be accepted as a part of the account. But Peacocke seems also to want to maintain that there is genuine explanatory content in the theory of grasp given in terms of implicit conceptions. In which case, the circularity involved must be minimised in some way so as to be unproblematic. How might this be?

After acknowledging the violation of the A(C) requirement, Peacocke writes:

Violations of the A(C) form are unobjectionable in the explication of a concept F because one can use one's own mastery of the concept F to assess what someone with an implicit conception involving F could be expected

to think or do in any given state of information. This is why a statement about what is involved in possession of a concept, and which does not respect the A(C) form, is not vacuous. It still makes an assessable claim. Each one of us, in evaluating the claim it makes, draws on his own mastery of the concept F being explicated. One draws on that mastery, and engages in simulations to assess what one would be obliged, or rational, to think or do in any given state of information. (2008: 145)

It is unclear how this provides any response to worries about the explanatory content of implicit conceptions theory. It might be that explanatory content comes from the fact that we, as theorists, can employ our own grasp of the sense. This would allow the view to be 'assessable' by us.

The shift to the theorist's perspective does not minimise the circularity so as to make it unproblematic. What we want is an explanation of what makes it rational to have certain inferential dispositions from the thinker's own perspective. The idea that circularity is not a problem because we can shift to the theorist's perspective, in specifying the content of mental states available to the thinker when performing certain inferences, means remaining silent on what needed explaining. In addition, the account would now presuppose that the *theorist* has an implicit conception. The circle is widened, but not in a way that lessens the problem (a never-ending chain of theorists explicating each other's concepts would not help). The idea that "one can use one's own mastery of the concept F to assess what someone with an implicit conception involving F could be expected to think or do" is of no help at all in figuring out *whether* someone possesses the concept (i.e. grasps the sense) in question. That is merely presupposed.

A different way the circularity might be minimised is to appeal to the idea that, as Peacocke says in the sketch quoted above, the inferences in which they figure are *sub-personal*. In various places, he characterises implicit conceptions themselves as sub-personal states. It might be that the circularity is present but is minimally problematic, because the inferential dispositions are rational and so personal-level, whereas the mental states that make them rational are sub-personal. Whatever exactly this means, it is not clear that Peacocke is entitled to say that implicit conceptions are sub-personal and maintain their explanatory role. In a comment on implicit conceptions, Burge writes:

I take the subpersonal level to be a level that is not only not conscious, but is not accessible to introspective or reflective consciousness and must be gotten at only theoretically. This is true of the basic grammatical structures underlying our linguistic competence and the information-processing structures underlying our perceptual experience. But... Peacocke takes implicit conceptions to be difficult but not impossible to make explicit through reflection. This makes it look as if implicit conceptions are really personal-level conceptions, just ones that are unconscious and relatively difficult to articulate in consciously available judgments. (Burge 2003: 384)

A fuller assessment would require working out exactly what the explanatory relation between implicit conceptions and rational inferences is meant to be. But however this might go, the idea that implicit conceptions are really sub-personal seems to be of dubious coherence, and so could not be what minimises the problem.

I conclude that a theory of grasp on the lines of implicit conceptions theory would involve the sense-theoretic explanation in circularity.

3.5.3 Diagnosis and generalisation

The fact that the problems for paradigms of the dispositional and rational basis approaches to the theory of grasp are so similar suggests that the source of the objections is something general about theories of grasp and how the sense-theoretic explanation of referential coordination can respond to the handles challenge.

Though they are ways of constructing senses that make use of different materials means that the precise nature of the problems are different, there is a pattern that can be discerned in both. Because it is part of an explanatory theory, a theory of grasp cannot mention thoughts that feature the sense in question. In other words, the theory needs to *re-describe* the phenomena. The theory changes the subject by saying what it is to have an ability to think thoughts featuring a particular sense in their contents without mentioning that sense. But then it is not clear why thoughts picked out in the new way are the right ones. In order to show that they are the right ones, the theory needs to show that thoughts of that sort are those that feature the sense in question. In other words, the theory needs to demonstrate how its redescription relates to what is being redescribed. The problem is that there are *multiple* ways of doing this, and there is no independent

reason for choosing one way rather than another. This means that the theory leaves open which thoughts have been picked out, the result being that the theory of grasp is not demonstrably adequate. The only way to determine which is the correct way is to presuppose the correctness of the theory, the result being that the explanation using that theory of grasp involves circularity.

The problems presented as CA1 and CA2 fit this pattern. Thoughts with senses in their contents are identified in terms of inferential dispositions. Idealised representation and rational explanation of those dispositions are ways of showing that the thoughts in question feature the relevant sense. Neither idealised representation nor rational explanation can be shown to fix the right dispositions without assuming the correctness of the theory of grasp in question. The problems presented as CM1 and CM2 can be traced to something else: generally speaking, if a theory can only be worked out by assuming that it is correct, then it is highly likely to be false, and the fact that theories are likely false makes it unsurprising that their truth needs to be assumed in order to motivate them.

The source of the pattern can be traced to two commitments of sense theory, one essential, and one highly unavoidable. The essential commitment is the individuation principle, which says that the individuation conditions for senses are to be given in terms of the thoughts and inferences that they make rational. The highly unavoidable commitment is that an adequate theory of grasp needs to employ the individuation conditions for a sense in order to ensure a one-to-one correspondence between a sense and conditions sufficient for grasping it. The thoughts and inferences that they make rational must be those that involve the sense for which one is giving a theory of grasp; this is true both in the general case, where one quantifies over senses in general, and in the specific case where one takes a given sense and specifies a theory of grasp for it. This means that it will be necessary for those thoughts and inferences in which the sense features to figure as part of the account.

This provides good grounds for thinking that the sorts of objections to the theories of grasp considered here will apply quite generally. If one must assume the correctness of a theory before doing anything with it, then it becomes unclear that one can really do anything at all, much less use it to give a substantive explanation. The sense-theoretic explanation seems to require assuming its correctness in order to motivate and formulate

an essential part of it, the theory of grasp. I conclude that it is safe to reject sense theory more generally as an adequate explanation of referential coordination.

3.6 *Conclusion*

Sense theory may seem to provide the perfect resources for giving an explanation of referential coordination. It posits and makes use of a kind of entity whose very nature is tied to patterns of rational inference. But it is really too good to be true. If there is nothing more to a sense than the patterns of rational inferences it is supposed to explain, then appeal to such things can have little explanatory content. This is illustrated by the failure of epistemic intensions theory and implicit conceptions theory to provide an adequate and non-circular theory of grasp. In the absence of any convincing reason to think that only sense theory will do, we have good reason to look instead at psychologistic explanations.

Chapter 4 - Typed Representations Theory

4.1 Introduction

This chapter examines the typed representations-theoretic explanation of referential coordination. Although typed representations theory can provide an alternative to sense theory, it faces a similar objection; there seems to be no way of filling out the notion of a type of mental representation that does not make the explanation fail to meet one or other of the explanatory desiderata.

Typed representations theory makes use of the idea that there are vehicles of mental content. These vehicles include mental representations at the level of complete thoughts (for example, the thought that Cicero is a Roman), and constituents of such representations (a representation of Cicero, and a representation of something's being Roman). The key idea is that these constituent representations can, in some way, be grouped together. This is often expressed by saying that they are of the same *type*. Advocates of mental representations have made appeal to such types in order to deal with Frege cases and associated problems without positing a more fine-grained sort of content. Rather than contents that determine reference, these theorists posit types of representations individuated in some way or other. This puts us here:

	Psychologism	Referentialism	Weight
Senses	No	No	Heavy
Typed representations	Yes	Yes	Heavy
Mental files	Yes	Yes/No	Heavy
Coordination functions	Yes	Yes	Light

There is room for convergence between sense and typed representations theories, but it is more interesting to look at typed representations theory understood as a referentialist and so genuine alternative to sense theory (§4.2).

Even as an alternative to sense theory, the typed representations explanation faces a problem very similar to the handles problem: it is one thing to say that there are types of representations, but it is another to say what it is that belonging to a type, their *type identity*, consists in. It is another thing again to provide an account of type identity that

provides a compelling explanation of referential coordination. Like the handles challenge, the *typing challenge* raises potential objections based on the adequacy and explanatory merits of the explanation that results from employing a way of typing mental representations (§4.3). The problem is clearly illustrated by a treatment of Frege cases given by Fodor in terms of the so-called *language of thought hypothesis*, making use of the idea of syntactic properties. The notion of a syntactic property is insubstantial enough that it can only be cashed out in a circular way that makes it devoid of explanatory content (§4.4).

More substantive proposals than that found in Fodor's discussion have been made, however. I consider two: Sainsbury and Tye's *originalist* theory, which types representations according to chains of deference which link back to an original introduction of the type (§4.5), and Prinz's *proxytype* theory, according to which representations are short-term constructions in working memory (§4.6). These proposals can be shown to be psychologically unrealistic or insufficiently personal-level. These problems are evidence for thinking that the type-identity of mental representations depends on rather than explains referential coordination, a conclusion reinforced by reflections on some remarks due to Millikan on what she mistakenly sees as the equivalence between what she calls typing rules and judgements of identity (§4.7). The conclusion is that psychologicistic theories need a richer set of resources than mental representations and their types.

4.2 *Typed Representations as an Alternative to Sense Theory*

This section clarifies the theory under examination, and shows how it offers a set of explanatory resources distinct from sense theory.

To begin with, what are typed representations? The representations in question are *mental* representations. Overlooking a number of complicating factors, a minimal gloss on what typed mental representations are can be given like so:

Something is a typed mental representation if it is (i) a mental entity that (ii) contributes its referential content to the content of a corresponding thought, and (iii) can be grouped together with other such entities in some way.

The ‘representations’ part of typed mental representations comes from (ii); the ‘typed’ part comes from (iii). The idea of a ‘type’, and its complementary notion of ‘token’, has its origins in the mechanics of the printing press, and is naturally associated with the particular way in which words, along with the letters that constitute them and the sentences which they in turn compose, can be grouped together. So, for example, the sentence ‘Every dog is a dog’ contains two tokens of the same type of word; we can say that they are *token distinct*, but *type identical*. Their type identity groups them together as tokens of the word ‘dog’. The application of this distinction to the idea of representations more generally is obvious. This includes mental representations, which are frequently, though not exclusively, cashed out as being very much like words and sentences. The image of ‘brain writing’ is frequently invoked for illustrative purposes, though it is important that mental representations be understood as psychological and not simply neurological in nature.

Though sometimes taken to be controversial (cf. Kaplan 1990), there is no need to worry too much about the metaphysics of types and their tokens here. For the purposes of giving a theory of referential coordination, a type is simply any way of grouping things together, and tokens are simply the things that are grouped. I will refer to a theory of what it is that groups mental representations together as a *typology*. Typologies are distinct from types; a type is a way of grouping things, a typology is a statement concerning what types there are. Typed representations theory can therefore be summarised as a commitment to the existence of entities meeting conditions (i-iii) in the gloss above, plus some such statement. The exact manner in which typed representations theory might provide an explanation of referential coordination will depend on the details of the typology, but the general idea is that the conditions that group representations together suffice to provide the indication of sameness of reference. A typology adequate to this explanatory task will be one that groups representations together at least partly in terms of properties that meet this condition.

Typed representations theory promises to avoid the problems with sense theory. A typology can make reference to a much broader range of factors than can the individuation conditions on senses. Mental representations have causal properties, and the properties needed to sustain these, and can interact with a thinker’s psychology in whatever forms that psychological processes can take. This means that typed representations theory can be much more closely anchored in the actual processes of thinking than sense theory.

As with sense theory, there is an initial worry to do with the fact that the particular theoretical commitments it makes are controversial. This is more true of mental representations than their being typed (if one grants the former then you get the latter for free). This is notwithstanding the presence throughout the history of philosophical theorising about the mind of appeals to mental representations of one kind or another at least since William of Ockham (see King 2007), to say nothing of Descartes, Locke, Hume and other luminaries, all of whom share the idea that mental representations present neat explanations for a range of facts about thought and other cognitive processes. Theorists often find it helpful to treat thought in terms of mental representations and their functional relations, such as the idea that beliefs are token representations in a ‘belief box’. It is not out of the question to think that the reason such treatments are helpful is because, in some way, they are *true*. At any rate, explanatory power is a large part of the argument for the existence of mental representations, and I take it that if they provide a good explanation of referential coordination, that would count as a strong bit of evidence in their favour. This initial worry can therefore be set aside.

A more immediate complication comes from the potential congruence between typed representations theory and sense theory. Although mental representations, as psychological entities, must be distinct from senses, as abstract entities, it is easy to overstate the difference between the two theories. This has been exploited by some defenders of sense theory who would also like to make appeal to the sorts of explanatory benefits they see as conferred by typed representations theory (see e.g. Peacocke 2005, Ichikawa & Jarvis 2013). The thinking here is not hard to make out: types of representations, as opposed to token representations, are a bit like senses, since they are abstract, and can occur in contexts (as words occur in sentences perhaps in conjunction with a context or other semantic features) so as to determine a particular referential content and hence a truth condition or proposition, or some other semantic complex. One might therefore want to say that the correct typology is one that types representations by the senses that they express—again, much as words and sentences are said to express senses.

This congruence does not warrant the conclusion that there could only be a merely ‘verbal dispute’ between sense theorists and typed representations theorists. The explanatory scope of typed representations theory is dependent on the typology that is

supplied; without a typology, it is a framework for explanation but will lack explanatory content. This need not be filled out with a sense-theoretic typology. The obvious problem for a typology that employs sense theory is that the handles challenge will simply arise anew. The rules of good housekeeping demand recognition of the congruence, but the exigencies of explanation favour developing typed representations theory as a genuine alternative.

Typed representations theory presents a genuine alternative to sense theory, at least, when the typology is referentialist; a typology is referentialist if it does not group together mental representations in terms of (and so does not impute to mental representations) properties that are more basic than referential content. In the remainder of the present study, I will therefore understand typed representations theory specifically as involving such a referentialist typology, and reference to typed representations should be understood as mental representations with a referentialist typology.

‘Referentialist’ is here being used in the sense introduced in §1.4, not in the sense employed by Edwards (2009) when discussing a version of typed representations theory advocated by Fodor to be examined in the next section. According to what Edwards calls ‘Concept Referentialism’, “concepts are individuated referentially. A constituent of a mental state is a token of WATER insofar as it bears the reference relation to water; and similar for DOG and the property of being a dog, BACHELOR and the property of being a bachelor, and so on.” (2009: 294) Types individuated by their reference will not be sufficient to give an explanation of referential coordination either: sameness of type *guarantees* sameness of reference, but sameness of type needs to provide an indication of sameness of reference, and there would obviously need to be more to the explanation than simply that they belong to the same type, which demands further resources. A typology that is referentialist in Edwards’ sense would therefore make the typing of mental representations effectively redundant.

4.3 *The Typing Challenge*

I said above that the initial worry about typed representations theory, that it involves controversial theoretical commitments, can be set aside in part because, if it offers a good explanation of referential coordination, then this would count significantly in its favour. There is, however, a general challenge that can be raised for that kind of explanation. The explanation of referential coordination by mental representations belonging to the same

type does not simply fall out of the notion of sameness of type. Since we are taking sameness of type as simply equivalent to being grouped together in some way, the way of grouping representations needs to be developed in such a way so as to show that sameness of type provides an indication of sameness of reference. We can put this issue like so:

Typing challenge: A typed representations-theoretic explanation of referential coordination is incomplete without, and so requires, an adequate typology.

Putting the challenge like this hopefully makes evident the similarity between this and the handles challenge for sense theory. Sense theory needs a way to say which senses are in the contents of, or can be used to type, which thoughts, so typed representations theory needs a way to say which types of representations are instantiated. That there is a similar problem is hardly surprising, given the similarity between senses and types of representations. In both cases, a condition on sameness and difference—of sense or type of representation—is needed so as to make the explanation complete. The potential for objections will be the same too. If there is no way of doing this without some explanatory problem creeping in, then the challenge indicates that the explanation is not a good one.

Nevertheless, the theories differ in more ways than they are similar, and the similarity once noted should not be taken as the end of the matter. Two differences are particularly relevant. One is that the range of possible typologies is going to be quite broad. Consequently, this chapter will look at a different typologies in order to make the case that the typing challenge presents more than an initial problem. There is no question of a diagnosis of a general difficulty, as was given in the last chapter to generalise the objection to sense theory, because of this breadth. But I think there is enough to indicate that typed representations are not the right way to go about explaining referential coordination. The other difference is that, since the potential objections that the typing challenge present will turn on what resources the typologies appeal to, and since these will differ from those of sense theory, the potential objections will differ also.

Whereas the entities posited by sense theory are guaranteed to be such as to cash out the basic idea that sameness of sense makes for referential coordination, this is not the case for types. A necessary condition on a typology meeting this end is that it must supply type identities that *match* patterns of referential coordination. Relations of sameness of

type must, at least, match relations of referential coordination between thoughts, otherwise it will not be plausible that sameness of type explains that feature of thoughts. There is a question about whether a proposed typology can be shown to meet this condition. The first potential objection is that, for any given typology, this cannot be shown. Call this the *matching objection*. The second potential objection follows on from the first: if the typology can be shown to meet this condition, does it do so in a way that provides an explanation without explanatory deficits? If not, then the explanation that results will be no good. Call this the *explanatory objection*.

4.4 *The Syntactic Typology*

The potential objections presented by the typing challenge are well illustrated by what is perhaps the most prominent version of typed representations theory, that owing to Fodor (1975, 2008). Fodor claims that Frege cases and associated problems can be explained in terms of the formal or *syntactic* properties of mental representations. Call this the *syntactic typology*. The problem with it is that the notion of a syntactic property is much too thin to avoid either the matching or the explanatory objections.

4.4.1 Mentalese syntax

Fodor's most recent treatment (Fodor 2008) is presented as a 'fresh start', superseding all previous efforts (op. cit.: 57), so I limit my discussion to that attempt. On Fodor's view, mental representations account for a collection of facts about thought, particularly the success of psychological explanation by way of attribution of belief and desire. The most important of these is the fact that psychology seems to make *essential use* of such mental states; they cannot be eliminated without serious and possibly total reduction in explanatory power. This feeds into two claims that provide impetus for typed representations theory: explanations involving intentional mental states are *causal* explanations, and the capacity to be in intentional mental states exhibits *productivity* and *systematicity*. Productivity is the fact that thinkers can be in an infinite or near infinite variety of different intentional states. Systematicity is the fact that there are non-arbitrary links between which intentional states a thinker can be in.

Fodor claims that these facts are best explained by the idea that thoughts consist of relations between thinkers and mental representations with a language-like structure. These representations have causal properties, and these causal properties are dependent

on those of the constituent mental representations (Field 1978 presents similar ideas). Constituent representations can be recombined in productive and systematic ways. In addition, Fodor claims that the causal processes that mental representations enter into are best understood as *computational* processes, akin to a Turing machine. In this combination of a representational and computational theory of cognition, the parallel with words and sentences touched on at various places above is made a feature of the proposal. Hence the explanatory hypothesis is known as the *language of thought* hypothesis, or *LOT* for short. The system of representations posited by LOT is sometimes referred to as *Mentalese*.

Fodor's development of LOT has largely taken place through showing opposition to alternatives to be inadequate. One alternative is a denial of the existence of intentional mental states, as in the eliminativist attitude to intentionality. Another is a denial of their theory-independent existence, as in the case of the instrumentalist position, notably defended by Dennett (1987), himself a trenchant critic of LOT (cf. Dennett 1991). Instrumentalism often comes with a *holistic* view of mental content, on which the smallest unit of mental content is a whole system of beliefs and desires and the inferential links between them. This means that the attribution of mental states to an individual (say, the belief that it is raining and the desire to not get wet) in order to provide an explanation of what she does (say, her grabbing an umbrella before going outside) brings with it a much wider system of beliefs, desires and inferences which characterises the entire mental life (or at least, the *rational* mental life) of that individual. A more moderate *molecuralism* contends that, at the very least, no belief or desire comes alone, but needs to come with a whole range of other similar states to be had at all; schematically, no belief about *a* that it is *F* can be possessed without having many other beliefs about *a* and other beliefs about *F* things.

Fodor places great stress in opposing both holism and molecuralism. This finds a clear expression in his *informational atomism*. On this view, the smallest unit of mental content is a lexical concept, or a word-like constituent mental representation. Debates about holism, molecuralism, and atomism mainly concern the metaphysics of intentionality. Accordingly, Fodor's atomism is complemented with an *informational* theory of intentional content on which intentional content is determined by nomic (law-like and counterfactual-supporting) relations between representations and what they represent (cf. Fodor 1987, 1990).

The strategy Fodor adopts to deal with Frege cases is one of divide and rule: distinguish different sorts of Frege cases, and show that LOT has the resources to account for them without needing to appeal to sense theory or a sense-theoretic typology (Fodor's reasons for preferring not to employ sense-theoretic resources are given at op. cit.: 52-7). He begins his discussion by distinguishing thoughts whose referential content is a function of representations that can be broken down into distinct parts, each of which has its own referential content (complex representations or 'concepts'), and those that cannot (primitive concepts). Fodor presents this as analogous to the *syntactic* differences between simple and complex noun phrases, as in the difference between the proper name 'Cicero' and the descriptive phrase 'the greatest Roman orator'. Syntactic differences can therefore distinguish concepts with the same referential content.

Syntactic differences depend on the formal properties of representations. Formal properties are those that, at least, account for the way in which representations enter into causal or computational processes. Since, as Fodor notes, "some concepts must be basic if any concepts are to be complex," (op. cit.: 66) syntactic properties must be extended to the primitive concepts. And, Fodor says, "[i]f there is a Frege problem, it must be about how to draw the type/token relation for (syntactically) primitive concepts. But if there is a Frege problem about primitive concepts, then it is resolved by appeal to their form, not by reference to their content." (op. cit.: 75)

4.4.2 Problems

Whether the syntactic typology avoids the two objections presented by the typing challenge depends, as Fodor suggests, on cases where it would need to be primitive rather than complex representations at issue. A problem is that the syntactic typology is not substantive enough to offer much help. This gets shown by a consideration of the 'Paderewski' variation of Frege cases, to which a large proportion of Fodor's discussion is dedicated. Recall the Paderewski case introduced by Kripke: this sort of case involves someone who gets presented with information about the same individual, with the same name, but for some reason does not understand the two occurrences of the name to refer to the same individual.

One might expect that LOT would have the implication that the structure of Paderewski cases can be as much a feature of cognitive as linguistic representation. The Paderewski-style case shows how the same word can be understood in different ways. LOT

representations are like words, so it might be the case that LOT representations of the same type can be ‘understood’ differently, in the sense that they enter into psychological processes differently. But if LOT types are individuated by syntactic properties, and if syntactic properties are whatever accounts for computational processes, then it follows that Paderewski-style cases cannot occur on the syntactic typology. Fodor (op. cit.: 72-3) proposes that we represent what is going on in such cases by means of subscripts; rather than distinct tokens of a *single* type, PADEREWSKI, there are distinct tokens of different types, PADEREWSKI₁ and PADEREWSKI₂.

But what makes for this difference between the types that the subscripts mark? Fodor’s answer is this:

I don’t care. *Type distinctions between tokens of primitive mental representations can be distinguished by anything at all, so long as the difference between them is of a kind to which mental processes are responsive.* Since, by definition, basic representations don’t have structures, type identities and differences among primitive Mentalese tokens are bedrock from the computational point of view. Tokens of primitive Mentalese formulas are of different types when they differ in the (presumably physical) properties to which mental processes are sensitive. (op. cit.: 79; emphasis added)

This is to give up exactly where an explanation is most required. Saying that syntactic properties distinguish types is insufficient, because the question is *which* of those properties distinguish types of the relevant sort if type differences match patterns of referential coordination. Saying that tokens are of the *same* type if they *share* syntactic properties is insufficient for the same reason. So the syntactic typology does not avoid the matching objection. Similar objections have been made before (cf. Ayedede 1998 for an objection concerning inter-personal typing of LOT tokens), but it does not seem that Fodor has much to say in response. We get no indication of what it might be that would make two tokens belong to different types other than: whatever it is that makes them belong to different types.

The matching objection aside, it is worth considering whether the syntactic typology could anyway really explain referential coordination. The syntactic typology is based on LOT, which is an account of the cognitive mechanism that accounts for the apparently causal nature of explanation by belief and desire, and the systematic nature of capacities

to be in those states. For a typology to explain referential coordination, it must be such that type identity provides an indication of sameness of reference that is adequately personal level. But syntactic properties are ill placed to do this, since they are paradigmatic examples of things belonging to sub-personal psychology. The only obvious way of avoiding this problem, or at least the only obvious way, is to define the relevant properties as those that explain referential coordination. This would clearly make the explanation viciously circular. That suggests that the reason Fodor's treatment of what distinguishes simple concepts gives out is that there is nowhere else to go without making the account manifestly inadequate. In any case, the syntactic typology does not avoid the explanatory objection either (Rives 2009: 222-4 presents a similar objection).

This second issue with the syntactic typology raises a general point about the prospects of typed representations explanations. Fodor's comment, quoted above, that type identities of LOT tokens are 'bedrock' suggests that he thinks any further questions about how LOT tokens work in psychological explanation is misplaced. But they are bedrock, as he also says, from the computational, or causal, point of view. The problem presented by Frege cases, with which Fodor explicitly engages, and the problem of referential coordination, with which he does so implicitly, are not simply a matter of causal goings on in the cognitive system. Their being bedrock from the computational point of view does not mean that type identities are bedrock from the *rational* point of view, whatever exactly that might mean. One cannot simply appeal to type identities and have done with it (similar remarks go for Heck's (2012) claim that formal relations between thoughts may be treated as psychologically primitive; Heck's focus is on explanation *by* intentional states rather than explanation *of* intentional states, and while his claim may be true in the former case, it is false in the latter).

Fodor seems to miss a fairly crucial point about the problems, one that should not be missed if typed representations are going to be of use in dealing with ways of thinking. What moved Frege was that someone might be rational in treating things with identical referential content differently. In the first instance, LOT is best understood as a hypothesis about the causal and dispositional goings on in the cognitive system. But a treatment of Frege cases needs to account for the rational properties of thought. Fodor seems to think that the resources of LOT can be extended to this second task, though his attempt by means of the syntactic typology is in the end rather half-hearted. The more general lesson is that a typology does not explain referential coordination, or any of the other problems of ways of thinking, if it is only in a position to explain the causal or

dispositional properties of thought. The resources available to the syntactic typology look not to be up to this task.

4.4.3 Summary

The syntactic typology provides a way of typing mental representations that is not substantial enough to match patterns of referential coordination (which syntactic properties are the relevant ones?) and seems to bear only on the dispositional rather than rational dimension of those patterns. A typed representations-theoretic explanation of referential coordination will need to go beyond typing mental representations in terms of syntactic or causal properties if it is to meet the typing challenge.

4.5 *The Originalist Typology*

Sainsbury and Tye (2012) have made a number of proposals framed as a way of dealing with a range of ‘problems of thought’, including Frege cases of several varieties, that provides an alternative to the syntactic typology. Their proposal individuates concepts in terms of *origins*, rather than anything analogous to formal, syntactic, or otherwise causal properties. Call this the *originalist* typology. The problem with this is that origins either fail to match patterns of referential coordination, or else require thinkers to have unrealistic cognitive capacities.

4.5.1 Origins and abstract continuants

Sainsbury and Tye endorse a number of claims made by Fodor in the context of propounding LOT, and recognise agreement with Fodor on the rejection of sense (op. cit.: 73), a referentialist view of content, a commitment to mental representations, as well as the idea that different causal properties may accompany representations with the same referential content (op. cit.: 85). However, where the syntactic typology Fodor endorses individuates types according to these causal properties, Sainsbury and Tye endorse a typology that individuates types according to what they call “originating use.” (op. cit.: 42) On this view, which they term *originalism*, mental representations should be understood as parts of non-eternal abstract continuants. Types are identified with these continuants, and like other continuants, such as ordinary objects, they can be distinguished by their origins and can undergo changes in their properties across time.

The substance of originalism is brought out by Sainsbury and Tye by exploiting the analogy between mental representations (which, as in Fodor's discussion, are called 'concepts') and words:

Words, like concepts, are non-eternal abstract continuants. They are invented or created, so there are times at which they did not exist.... Words are not individuated by their spelling, since the same word can be spelled in different ways (e.g. UK "colour", US "color"). Words are not individuated by their pronunciation, for the same word can be pronounced in different ways (e.g. "lieutenant"). Words are not individuated by some combination of spelling, pronunciation and reference, for there might have been two orthographically and phonetically indistinguishable names for the same thing.... Rather, words are individuated by their origin. (op. cit.: 42)

The typology for mental representations is explicitly stated later, exploiting this analogy:

Both words and concepts invite the same challenge: what makes a use of a word or concept a use of one word or concept rather than another? The answers take the same form in both cases: there are originating acts in which words or concepts are introduced, and these form the basis for subsequent propagation through the linguistic or conceptual community. Non-originating uses are actions that are in some way dependent upon earlier uses, by the same or other users. (op.cit.: 59)

Although uses of mental representations are said to be "concrete manifestations," (op.cit.: 64) it is left slightly unclear what the tokens are, though one can assume that they are things like LOT tokens. The important point (to reiterate something said above) is that, while LOT tokens have syntactic properties that explain their role in computational processes, those properties do not supply the only typology or the typology that will provide the best explanation of various aspects thought. Sainsbury and Tye's alternative is to type representations in terms of the abstract continuants to which they belong.

Does the originalist typology avoid the troubles of the syntactic typology? Much turns on how the relation of dependence tracing back to on original use, which is what constitutes

a token belonging to a type, should be understood, and a problem arises from the analogy with words on this score. The problem can be finessed by considering the matching objection. The idea of original use taken at face value may suggest that sameness of type individuated by original use does not need to, nor seem very likely to, match patterns of referential coordination between thoughts: two tokens might be part of the same abstract continuant, though the thinker does not treat them as referring to the same thing. This seems to rule it out as an adequate typology. But the face value of the idea of original use may not be the correct one. There is more to be said here, though seeing what exactly means going into how the originalist typology fares with the typing challenge.

4.5.2 Problems

As with the syntactic typology, it is useful to look at what is said in the context of Paderewski-style cases to see how and if the proposed typology works. There is some difficulty in working out what exactly the view is meant to be. There is the official presentation, which seems to make type identities fail to match patterns of referential coordination, and there is a slightly amended presentation which makes the appeal to originalist types psychologically unrealistic.

On their official presentation (op. cit.: 131-8), Sainsbury and Tye claim that originalism implies that Paderewski cases are those where two tokens are of the same type. This contrasts with Fodor's view that they are of different types (as they recognise, op. cit.: 133n9). As a consequence, Sainsbury and Tye are led to the view that such cases involve a thinker with false, albeit rational, beliefs about their concepts (op. cit.: 135), or a second-order belief about their beliefs that turns out to be mistaken, to the effect that they are about different things (op.cit.: 137). There is nothing obviously wrong about this, though positing such second-order beliefs is not very neat. The problem is that it makes the appeal to originalist types redundant. Sainsbury and Tye implicitly recognise this when they say "Paderewski examples are consonant with originalism, but finding a suitable description of the cases makes little use of details of the theory." (op.cit.: 138)

Given the official treatment of Paderewski-style cases, Sainsbury and Tye seem committed to a typology that hamstring the typed representations-theoretic explanation. However, it should be noted that, though the analogy with words pushed down the originalist rather than syntactic path provides a significantly more substantive typology, the crucial

part—how particular uses get to have membership of the same abstract continuant—remains under-specified. Type membership means having the same originating use, but what is to have the same originating use? It is surprising that Sainsbury and Tye's direct comments on this are highly schematic (one is reminded of Fodor's reticence when distinguishing types of primitive concepts). They say that uses are 'R-linked' (op. cit.: 44, 86), and that the relation R is that of 'descent' (op. cit.: 44), though not much detail is given on the nature of R. After a short overview of relations between uses of words, we are told that

[t]here is a concept-originating event, an act of subject S_1 , in which concept C is originated. This generates a C-reproducing mechanism in S_1 , which can create copies or tokens of C. Being produced by this mechanism is what makes S_1 's later use of a concept a use of C. (op. cit.: 60)

Perhaps the idea of a 'C-reproducing mechanism' gets us a bit closer. But mechanisms of this sort are characterised only in terms of what they achieve, the reproduction of C, and we wanted to know what this was, so this tells us very little.

More of a grip on what Sainsbury and Tye have in mind is provided by their idea that the non-eternal abstract continuants that constitute types of mental representations are *use-trees* (op. cit.: 88). These trees are sets of relations of R-linking, which Sainsbury and Tye call 'deference involving' (op. cit.: 44), and uses, which as concrete things can be identified independently of their types. The obvious suggestion is that referentially coordinated thoughts are those employing uses belonging to the same use tree. This indicates that the matching objection can be avoided, and that their official presentation of the Paderewski case, and the face value understanding of original use, is not essential to the originalist typology. The originalist need not say that Paderewski cases involve two tokens of the same type. If deference is understood as treating as referring to the same, then it looks like the thinker is *not* deferring in her use of two token representations in such a case. So it must be the case that there are two distinct use trees in such cases, hence two distinct types.

However, once the notion of deference is given this role in the originalist typology, and the abstract continuants are specified as use trees, it becomes clear that the matching objection is avoided only at the cost of inviting the explanatory objection. Characteristically, what deference involves is left underspecified. We are told that

deference takes the form of intending to use the concept as it has been used by oneself or others on previous occasions. (op. cit.: 42)

Later, they write that

[d]eference can be modelled (rather over-intellectually) as the recognition that others already use a concept, together with the desire to use the very concept they use, with the very reference it has in their uses. (op. cit.: 70)

Taking seriously the idea that deference, whatever else it is, involves intentions to use a mental representation in a particular way, the problem is that this attributes cognitive capacities to thinkers that we have no reason to think they have. The rather over-intellectual model of deference is not something that can be avoided; how else might it be cashed out? *Any* way of cashing out of deference, at least any that is not simply defined as whatever can explain referential coordination (or the various kinds of Frege case), will presumably need to attribute such capacities that involve recognition of uses of concepts and intentions to use concepts in the same way, and this level of reflective access is not something we have any reason to attribute to thinkers independently of the theory.

The originalist typology therefore faces a dilemma: if types are abstract continuants not linked by deference, then sameness of types does not match patterns of referential coordination, and so originalism would face the matching objection; if types are abstract continuants that are linked by deference, then sameness of type perhaps does match relations of referential coordination, but the explanation of this fact is psychologically unrealistic, and so originalism would face the explanatory objection.

4.5.3 Summary

The fact that the key notion of deference is dependent on thinkers recognising and desiring to use their mental representations in particular ways, along with Sainsbury and Tye's official endorsement of the same-type rather than different-type view of Paderewski cases, suggests that the originalist typology is really too closely modelled on how words in public language work (compare Millikan's criticisms of originalism in her 2011). This is another demonstration that not any typology will do, and that typed representations

theory may not be as good an alternative to sense theory as it promised to be. A more radical departure seems called for.

4.6 *The Neo-Empiricist Typology*

The typologies examined in the last two sections both take their cue from an analogy between mental representation and words; mental representations are thought of as *sentential* (or quasi-sentential). Prinz (2002) has proposed a theory of mental representations that he terms *neo-empiricist* that stands in contrast to this. It shares the classical empiricist idea of cognition as depending on perceptual experience, backed up by appeal to developments in recent cognitive psychology, particularly the work of Barsalou (1999). On Prinz's view, types can be individuated by the properties encoded by representations constructed in short-term working memory out of perceptual representations stored in long-term memory. Call this the *neo-empiricist* typology. The problem with this is that, although differing from the originalist typology in obvious ways, it has essentially the same pattern of problems.

4.6.1 Stored perceptual representations

One way to express the disagreement between the neo-empiricist and the sententialist is over the question of whether mental representation is *amodal* or *modality-specific*. The modalities in question here are perceptual modalities, such as visual, aural, haptic, and so on. On the sententialist view, a given mental representation is not tied to any particular perceptual modality; it has its representational properties independently of any perceptual inputs into the cognitive system. On the empiricist or neo-empiricist view, by contrast, mental representations are “copies of perceptual representations” (Prinz 2002: 108), and so “are couched in representational codes that are specific to our perceptual systems;” (op. cit.: 119) they have their representational properties only in virtue of the deliverances of particular modalities.

Whether or not neo-empiricism in its extreme form is defensible, one might well think that the richer resources of modality-specific representations could avoid the problems with the syntactic and originalist typologies. The basis of neo-empiricism in the empirical psychological literature certainly lends it a significant line of defence against the charge of being psychologically unrealistic, and the problems with the sententialist typologies arose

due to the difficulties of cashing out an account of type-identity made in the image of word types suitable for cognition.

In contrast to the traditional empiricist appeal to mental images, Prinz argues that mental representations employed in cognition are *proxytypes*, so called because they “stand in as proxies for the categories they represent.” (op. cit.: 149). Proxytypes are constructions in short-term working memory out of bundles of perceptual representations stored in long term memory across many modalities that encode properties in modally specific ways (op. cit.: 148-152). The properties encoded in the proxytype account for the categorisation and identification behaviour associated with a particular kind of thought (op. cit.: 161-3). So, for example, thinking about dogs will lead one to think about furry animals because one’s perceptual representations of dogs involve the feel of fur and the visual look of a medium-sized animal.

Prinz applies his theory to what he calls ‘cognitive content’ (see Prinz’s list of desiderata given in §2.4), though what exactly he means by this is somewhat unclear. He elaborates it in terms of Frege cases and Twin earth cases (see op. cit.: 270), and states outright that “two people have the same cognitive content, on this proposal, when they have type-identical proxytypes.” (ibid.) This suggests that Prinz is operating with a typology that he takes to be sufficient for explaining cases of these kinds. But how are proxytypes typed?

Along the lines of Fodor’s primitive/complex concepts distinction, Prinz notes that a typology needs to provide “a set of primitives that are not type-identified by further features.” (op. cit.: 273) He says that there are two ways to type proxytypes: one is in terms of their referential content, or what they represent, which Prinz calls a proxytype’s *real content*; the other is in terms of the properties that they encode, which Prinz calls *nominal content* (the nominal/real contrast is meant to echo Locke, cf. op. cit.: 277). Prinz proposes that proxytypes can be typed in terms of either kind of content:

A pair of concept tokens can be identified in virtue of sharing their real contents or in virtue of sharing their nominal contents. Call these “real types” and “nominal types.” Any given token has both a real type and a nominal type. (op. cit.: 279)

Sameness of either content is sufficient for sameness of type; though note that these are distinct typologies. A dual typology of this kind is available to any typed representations

theorist, though the appeal to nominal types is distinctive because of its reliance on properties encoded by stored perceptual representations. The appeal to nominal types is what is distinctive about the neo-empiricist typology.

4.6.2 Problems

Does the neo-empiricist typology fare better than the sententialist typologies? Not really. On Prinz's official presentation, nominal types fail to match patterns of referential coordination. Bringing in the only obvious amendment makes the appeal to nominal types either circular or unrealistic.

As before, a test of the neo-empiricist typology is in how it deals with a Paderewski-style case. Prinz discusses a case of "a person who fails to realise that Farrakhan the violinist is Farrakhan the religious leader." (op.cit.: 271) What he says there is not much more satisfying than what we have seen up to now. Prinz recognises that this case "can arise where no difference exists in the perceptual representations constituting our proxytypes," and

[H]aving two distinct Farrakhan representations might lead one to say, "Farrakhan is not Farrakhan." The point is that any difference in Farrakhan-directed behaviours is a function of this belief in the existence of two Farrakhans, not a function of anything intrinsic to the two Farrakhan representations.... Consequently, these two representations can be said to have the *same cognitive contents*. A person in this situation has *distinct representations* in long-term memory (think of two mental file folders without any contents), but *their cognitive contents are the same*. (ibid., emphasises added)

As with the official presentation of the originalist typology, the problem here is that Prinz's take on the case looks to make sameness and difference in types irrelevant, as he acknowledges: "In this special case, the fact that one can be surprised to learn that Farrakhan is Farrakhan is explained by *something other than cognitive contents*." (ibid.) The similarity with the originalist treatment of this sort of case is striking; there, we needed to amend the official presentation so that the appeal to types was not redundant, and the same goes here. The neo-empiricist typology looks to be vulnerable to the matching objection. Is there a fix?

Prinz wants to say that this case is one where there are two representations with the same cognitive content, even though the thinker does not treat them as referring to the same thing. But that would seem to be a paradigm instance of difference in cognitive content, given that cognitive content is what cuts finer than referential content and can explain patterns of rational inference, which is what is at issue in Paderewski-style cases. It is possible that Prinz's cognitive content is really *descriptive* content. The contrast, in his terms, is between intentional content, which is what a concept represents, and cognitive content, which is the description that concept associates with what it represents. If this is right, then it is not clear that nominal types will help with explaining referential coordination. If nominal types are individuated by descriptive content, then there seems no way to get descriptions to individuate types in a way that matches patterns of referential coordination without inviting the explanatory objection.

This worry can be seen by applying the proposal to the Farrakhan case. We might amend Prinz's official presentation of nominal types so that there *is* a difference in cognitive content in this case. In which case, there needs to be a difference maker. And as Prinz says, the difference could be the fact that the thinker has the belief that there are two people who happen to have the same name, rather than one. That would be information that distinguishes two things. We can call this *distinctness information*. An amended neo-empiricist typology includes distinctness information to distinguish otherwise identical types of mental representations.

The problem is that including distinctness information in the typology makes the account either circular or psychologically unrealistic. A typed representations explanation is committed to there being two representations with the same referential content, but of different types. Difference in type is being explained by the presence of a belief. Either that belief involves representations of these types, or it does not. If it does, we have a circle; a belief that two things are distinct is being made use of in the explanation of what makes it possible to believe that they are distinct. One would need to employ those type-distinct proxytypes, as part of an explanation of what makes them type-distinct. If it does not, then the account becomes unrealistic. What belief would do? Presumably, only a belief *about one's proxytypes*, that they are representations of distinct things. In which case, the explanation attributes to thinkers beliefs about their stored perceptual representations. This is psychologically unrealistic (saying that this might be 'implicit' belief is no help; recall a similar problem with Peacocke's implicit conceptions theory).

Therefore the account avoids the matching problem only by inviting the explanatory objection.

Prinz says the Farrakhan case is ‘special’, and so may wish to say that it does not undermine the neo-empiricist typology. This might be right in the sense that it is difficult to think of a plausible case where one has no information whatsoever about two distinct objects, other than that they are distinct. But there can be closely related cases where a thinker has a quite a lot of information about two things, but no information that distinguishes them. This would seem to be a lot less special and a lot more common. Besides, if the amended neo-empiricist typology is needed to work in the special case, then it will always be needed; the peculiar character of the case only strips things back to show up the problem clearly.

The Farrakhan case shows that the neo-empiricist typology faces the explanatory objection because it needs to distinguish types by reference to those types, or else by attributing to thinkers unrealistic levels of insight into the nature of their mental representations. The need for an amendment, and so this pattern of problems, extends to cases where there *are* differences in encoded property as well. Take the following instance of the paradigmatic conjoin-and-generalise inference:

- (1) Earlier, at t , Lucy was wearing a hat.
- (2) Now, at $t+1$, Lucy is not wearing a hat.
- (3) Therefore, there is something that was wearing a hat then and is not wearing a hat now.

How would the neo-empiricist typology explain a thinker making this inference? On a non-sentential typology, there is no distinction in the structure of the mental representation between the Lucy-representing part and the hat wearing-representing part. The first two thoughts would need to involve type-identical proxytypes referring to Lucy. But in the first, she is represented as a hat-wearer. In the second, she is not. So the proxytypes in each case would encode different properties, and so would be type-distinct. In which case, it should not be possible for the thinker to rationally make this inference, if sameness in type is what explains patterns of rational inference. And yet the inference is clearly one that could be a rational one to make. In which case, the neo-empiricist typology does not match the patterns of referential coordination. An amendment is needed; in this instance, information that Lucy wearing a hat at t is the same as Lucy not

wearing a hat at $t+1$. Again, the problem is that this either puts us in a circle, or makes the account unrealistic.

It is striking that, though covering thoughts about a number of 'abstracta', such as democracy, causation, and disjunction (Prinz 2002: 165-188), Prinz dedicates no space to a discussion of how time might be represented. Internal *awareness* of change does play a role in his account of thoughts concerning causation, essentially an update of a Humean account with the resources of proxytype theory, making use of the awareness of past immediately successive events and expectation of future immediately successive events (Prinz 2002: 173-177). But we are not told exactly how a thinker is to represent one and the same object as undergoing a change over a single period of time and so what would make it possible to take anything from an experience of successive events. Given the importance of causation, change, and time to our conception of things, the fact that they take up a blindspot of the neo-empiricist typology is a problem, and the difficulty for the explanation of referential coordination simply makes it acute.

Prinz may want to reply that the neo-empiricist typology is directed at something other than what I am insisting presents a problem for him. The explanation of referential coordination involves the rationality of inferential dispositions. Proxytype theory and the neo-empiricist typology that it supports is informed by investigations into information-processing mechanisms undertaken by cognitive scientists. This mechanism operates according to the dispositions that are grounded in the various components, representational and operational, that make it up. Proxytype theory, it might be said, is a theory of this mechanism, and so is only bound to give an explanation of how the dispositions are grounded. The neo-empiricist typology is a way of accounting for certain patterns of these dispositions.

This reply concedes the point that the neo-empiricist typology cannot explain referential coordination. It just tacks on the rider that it is not *supposed* to, which is not much help. But it is not clear that Prinz himself is in a position to make this sort of concession. His typology is meant to account for cognitive content, and cognitive content is what is supposed to be at issue in Frege cases. That aside, the aptness of the reply can be granted. But it raises perhaps the deepest problem with the neo-empiricist typology as an explanation of referential coordination: it is unclear how a nominal type, individuated by the properties encoded in stored perceptual representations, could by itself provide an

adequately personal-level explanation. This is an instance of the general problem I flagged at the end of §4.4.

What the typology *can* provide, perhaps, is an explanation of the information-processing dispositions possessed by a cognitive mechanism, of the kind at issue in psychological theories of concepts. But we are looking for an explanation of the reasons a thinker has for being prepared to perform a certain kind of inference, and this is a distinct issue. If the neo-empiricist typology is to work as explanation of referential coordination, they need to be connected. If they are connected, we need some sort of story about how the former bears on the latter. And neither the neo-empiricist typology, nor proxytype theory more generally, provides resources for doing this. There is no contact with what a thinker has reason to do, only with the operations of a cognitive mechanism. Even supposing the matching objection could be overcome, and the difficulties of circularity and psychological realism could be avoided, whatever indication of sameness of reference is provided by sameness of nominal type is not one that can function as a *reason*.

After effectively admitting that the neo-empiricist typology is not sufficient for explaining what he calls cognitive content, Prinz adds the parenthetical instruction to “think of two mental file folders without any contents.” (ibid.) This remark is made in passing, but should be taken seriously. Prinz’s suggestion is effectively that what provides the indication of sameness of reference concerns how information is stored or processed, rather than some way of typing representations that can be given independently of this. Mental files are entities that have a similar role in a cognitive system to proxytypes, though they involve making far fewer commitments as to the format of mental representation. Coming up with a way of linking information processing dispositions to the reasons a thinker has to perform inferences should go via an account of these entities rather than the comparatively robust long-term memory networks and proxytypes, or indeed any other way of typing mental representations. Doing this means examining mental file theory, which will be looked at in the next chapter.

4.6.3 Summary

The neo-empiricist typology involves a comparatively rich way of individuating types, but even so, it does not avoid the problems raised by the typing challenge. The conclusion that we could draw is that typologies of mental representations that match patterns of referential coordination do so because they *follow* those patterns, and not the

other way around. Consequently, they will not *explain* patterns of referential coordination. The kinds of properties of mental representations that one could give independently of these patterns will not supply a typology capable of the explanatory work required.

4.7 *Sameness Markers*

Before moving on to consider the kind of information-processing structure provided by mental file theory, I want to consider some related claims made by Millikan (2000) in the course of her treatment of what she terms *grasping sameness* and *sameness marking*. Millikan's views, and no less her expression of them, are somewhat complex, involving a number of sometimes subtly distinct formulations. They do not fit neatly into any of the three theoretical approaches, but can rather be seen as advocating the idea that, to explain referential coordination, one needs to look at it functionally, abstracting away from how it is implemented. This idea is of a piece with the call to look at the properties of the information processing mechanism. However, her view on what this involves is insufficiently sensitive to the fact that sameness of reference figures in thought as a rational phenomenon.

4.7.1 Typing rules and identity sentences

To begin with, what is sameness marking and grasping sameness? Sameness is grasped when a thinker 're-identifies' (or 'coidentifies') something, which Millikan glosses at one point like so:

For a perceiver or cognizer to re-identify something JUST IS to be disposed, or for some subsystem of theirs to be disposed, to pair representations of that thing in perception and/or thought as a middle term for mediate inference, or other amplificatory information-processing, and/or for guiding action. (op. cit.: 144; emphasis in original)

Sameness marking is what is *used* to grasp sameness:

What makes a marker a sameness marker is that the perceptual/cognitive system use it to control the mediate inferences and other content pairings

that they make in guiding amplificatory information-processing an action.
(*ibid.*)

Putting aside the involvement of perception, re-identification can be understood as effectively equivalent to what I have been calling referential coordination. Mediate inferences are those that involve putting together thoughts in inference. Sameness markers can therefore be understood as what provides the indication of sameness of reference. It will become apparent that this interpretation needs qualification, so I will stick with Millikan's terminology for the time being.

Much of Millikan's argument is dedicated to showing the inadequacy of Fregean accounts of mental content (*op. cit.*: 159-176). Her argument against Fregean accounts of mental content begins with a discussion of what she calls *models* of sameness marking, how the mind or brain (*cf.* her revealing remark at 2000: 140) marks sameness, beginning with what Millikan variously terms the 'Strawson model', 'dot markers', or the 'naive Strawson-model' on which "all those individuals [a thinker] knows of are represented by dots, and the predicates the man knows to apply to each are written on lines emanating from these dots..." (*op. cit.*: 136) (these dot markers can be recognised as basic mental files, see §5.2). Another model is the duplicates model, that "pictures thoughts each as a separate sentence token in a mental language," (*op. cit.*: 137), familiar from the discussion of typed representations. She also mentions an 'equals sign model', on which "a mental equals sign" uses or "rides piggyback on the duplicates marker, indicating *examples* of two different duplicatable types," the effect of which is "that all tokens of either exemplified type are then treated as representing the same." (*op. cit.*: 137-8). Along with these, she mentions a 'synchrony model', a 'Christmas lights model', and an 'anaphor model', which we need not discuss.

The crucial claim that Millikan makes is that the models of sameness marking are functionally equivalent to dot markers (*op. cit.*: 160). This means that they involve the brain or mind using something in just the same way as it would a dot marker; what matters is the *use*, not the choice of model. Millikan draws from this the conclusion that Fregean treatments of mental content are mistaken (*op. cit.*: 168-70). This argument presents significant interpretative difficulties (*cf.* Lawlor 2005), and can be set aside.

More important is that, during the discussion of why the functional equivalence of all models of sameness marking to dot markers has this consequence, Millikan remarks that

the only exception to the functional equivalence of sameness marking to dot markers is the equals sign model (op. cit.: 160). Millikan proposes "to explore this model... in its most naked form, namely, that in which discursive thinking is analogised to the unfolding of a formal system." (ibid.) In such a system, 'identity sentences' of the form

$$a = b$$

are the *same* as what she calls 'typing rules' governing the syntax for the language (op. cit.: 162). When the terms for composing well-formed formulas are given, the individual constants are set out as letters, with a rule that all letters of the same type receive the same interpretation (op. cit.: 165). This allows for trivial transformations, such as

$$F(a) \dashv\vdash F(a)$$

Introducing an identity statement allows one to make non-trivial transformations such as

$$F(a) \dashv\vdash F(b)$$

So, as far as a formal system is concerned, identity statements are logically indistinguishable from typing rules. They have no distinct role to play in characterising a logical language, the only difference being that typing rules are needed to prevent regresses in formulation, for long familiar reasons (Carroll 1895).

Similarly, Millikan claims, identity *judgements* are functionally the same as typing rules for mental representations:

For the mind, there... is no distinction like that between an identity axiom or postulate, $A = B$, written at the top of the page, and a typing rule. For there is no distinction like that between what is written on the paper and what is written in the structure of the reader—in the structure responsible for conforming the reader's reactions to a certain typing rule.... Write an identity sentence, that is, a structure responsible for producing certain coidentifications, in neuronal patterns instead of in graphite, and the distinction between identity sentence and interpretation mechanism vanishes.... Marking sameness, however that's done, and fixing identity beliefs is exactly the same thing. (op. cit.: 167)

She goes on to conclude that, “as distinguished from an identity sentence or assertion, there is no such thing as an identity *judgement*.” (op. cit.: 172) Although she does not make it fully explicit, the conclusion is that the non-equivalence of the equals sign model to the Strawson model does not matter because it requires identity judgements to play a distinctive role, which they do not.

4.7.2 Millikan’s mistake

Millikan’s discussion raises a couple of points. One is that typed representations offer only an implementation of a general functional aspect of cognition (‘grasping sameness’), and that a proper understanding of that function can abstract away to a certain extent from how it is implemented. This point seems correct, and to some extent, the positive account to be offered below will be in this spirit. The other is that we can understand that function by means of seeing what follows from an implementation of the Strawson model, and dispense with identity judgements. This point is not correct. It involves mistaking a rational phenomenon for a merely dispositional one. Grasping sameness cannot simply be the use of dot markers and their functional equivalents, for this reason.

Providing *some* way of typing representation that is many-one related to referential content is not enough. There needs to be some explanatory content provided by the typology, and having this explanatory content involved meeting the condition that it explain why having the paradigmatic inferential dispositions would be rational. By contrast, Millikan does not distinguish between typologies in this way, at least within the scope of thinking about what it is for a mind or brain to grasp sameness. It is precisely her point to instead claim that, functionally speaking, they are all the same—that is, they are all functionally equivalent to dot markers. Grasping sameness, for Millikan, is therefore a matter of having a certain set of dispositions towards one’s information, old and new. This does not allow for the fact that grasping sameness concerns the *reasons* thinkers have for organising their information in the way distinctive of treating them as being about the same thing. This exerts an explanatory demand on theories of grasping sameness that naïve Strawson markers and their functional equivalents cannot meet; or, perhaps, if they can then this remains to be demonstrated, and we have reasons to think they cannot.

The dispositions on which Millikan focuses are involved with recognitional abilities, or abilities to re-identify, which she terms ‘substance concepts’ (an unhelpful choice of words, as Millikan has recognised, cf Millikan 2013: 281). Strictly speaking, Millikan denies that abilities are simply dispositional (see her discussion at 2000: 51-63), so it is possible to object that she is really only concerned with these. One might also point out that the problem cannot be exactly that the functionally equivalent ways of grasping sameness are inadequately personal level, which was one of the main problems faced by typed representations theory, since they seem to be functions exercised at that level. The point still stands that her discussion neglects how the abilities to re-identify, and the way that sameness markers structure these abilities, have a rational aspect; there are reasons that bear on how they are exercised, and naive Strawson markers by themselves provide no account of how this can be.

Another way to come at the problem here is by considering the point that, while the obvious way for identity to figure in thought is as the content of a judgement to the effect that one thing is identical to another, this cannot be how referential coordination works. Apart from the fact that this merely introduces another step in the inference in need of coordination, there is also the plain fact that identity judgements (of the informative variety) *see past* differences in ways of thinking, which must be different from whatever sameness in way of thinking might be. The point is that identity judgements and whatever accounts for sameness in way of thinking have distinct roles in our cognitive lives. Identity judgements can be subject to further inquiry, to be confirmed or disconfirmed, and the consequences of either would be of a different kind entirely from continuing or dissolving a single way of thinking. Identity judgements, therefore, cannot be functionally equivalent to typing rules that provide types that reflect relations of referential coordination. They have different rational—hence functional—roles.

This difference between identity judgements and referential coordination is seemingly invisible to Millikan because she neglects the fact that grasping sameness, and recognitional and re-identificatory abilities, are rational phenomena.

In light of the fact that Millikan thinks that her claims about sameness marking speak against Fregean theories of mental content, it is worth returning briefly to the topic of how typed representations theory figures as an alternative to sense theory. Those aligning themselves with Frege's reflections on the cognitive value of identity statements, and his broader concern with the nature of inference, are often primarily concerned about

questions to do with rationality (Burge 1993a, Peacocke 2005). As we have seen, and in sympathy with Frege's general lack of concern with the actual psychology of thinking, Fregeans tend to be not so concerned about the actual process of thinking, except insofar as this helps articulate what it is to grasp a sense. Those opposed to the Fregean tradition tend to be less concerned about rationality, and more foundational questions about the psychology of intentional mental states, particularly the basis of the capacity to have thoughts and the explanation of the properties of that capacity. Hence the criticisms levied by opponents of Fregean sense are often unpersuasive, their treatments of Frege cases unconvincing, and their views are open to being simply assimilated to the Fregean view (that both sides often put their claims in terms of concepts only makes matters worse).

It is possible to think that the idea that there is something to do with rational inference that is more fine-grained than reference is sufficient for a Fregean view of thought. But this is unhelpful at best. It overlooks what is distinctive and substantive about Frege's contribution, and what many have found appealing in it. As I have argued, this contribution is ultimately not going to prove satisfactory when extended to explaining referential coordination. And that may lead one to think that, if there are problems with the Fregean notion of sense, then one can disregard rationality altogether, and simply concentrate on dispositions. Millikan's arguments against Fregean theories of mental content can be seen in this way. She rejects modes of presentation—Fregean senses—and so is required to deny a distinctive role to identity judgements. The mistake on the latter point indicates a mistake on the former, in argument if not in conclusion.

It is possible both to reject the appeal to sense theory and to allow thought to be more than just something that simply happens to us. The problem then is to work out a more satisfactory alternative than typed representations theory provides. To follow up on Millikan's suggestion to concentrate on the functional aspects of grasping sameness, it is necessary to look at how information is stored and processed in inference, with an eye to seeing what form of rationality this might supply.

4.7.3 Summary

The problems with typed representations theory might make one think that the right way to go is to abstract away from particular implementations, so avoiding the need to provide substantive typologies, and instead concentrate on the functional aspects of

referential coordination. Millikan makes this point, but takes a one-sided view of what that would involve, a view that leads her to overlook the important rational differences between inferences that involve identity judgement and those that do not. Preserving this difference means taking seriously the rational dimension of referential coordination. Doing so does not mandate a Fregean approach so long as a more satisfactory psychologicistic approach can be found.

4.8 *Conclusion*

Typed representations theory provides an alternative to sense theory, but it fails to offer a significantly better explanation of referential coordination. A typology is needed that both matches patterns of referential coordination and provides an indication of sameness of reference of the appropriate sort to do the explanatory work. Neither the syntactic typology proposed by Fodor, the originalist typology proposed by Tye and Sainsbury, nor the neo-empiricist typology proposed by Prinz negotiate the typing problem without falling short on either of these points. These problems need not deter one from pursuing the suggestion that a psychologicistic explanation will do better than sense theory, but if it can, then a different conception of the psychological entities and properties involved is needed. Millikan's suggestion to view typed representations in terms of the information processing functions they implement points to a means of doing this, but her dismissal of identity judgements as playing a distinctive role in inference shows that it needs to be sufficiently sensitive to the rational aspect of how identity figures in thought.

Chapter 5 - Mental File Theory

5.1 Introduction

This chapter evaluates mental file-theoretic explanations of referential coordination. Mental file theory overlaps with both sense theory and typed representations theory in some respects, but deserves independent consideration as it makes available explanatory resources that are inessential to or incompatible with those theories. I will examine some treatments of mental file theory in the literature, and argue that they fail to provide explanations of referential coordination that are both psychologically realistic and sufficiently general.

The exact nature of explanation by mental files is relatively up for grabs. Mental files are psychological entities, and existing treatments of mental file theory tend to tie them closely to the explanation of referential content, and so can be considered heavyweight. Whether this explanation of referential content involves a more fine-grained notion of content depends on particular applications. This puts us here:

	Psychologism	Referentialism	Weight
Senses	No	No	Heavy
Typed representations	Yes	Yes	Heavy
Mental files	Yes	Yes/No	Heavy
Coordination functions	Yes	Yes	Light

As a way in to seeing how explanation by mental file should go, I consider two worries that might be raised about any such explanation: that talk of mental files cannot, or should not, be understood literally, and that they can provide only vacuous explanations. There are sound reasons to reject both worries. Mental file theory can be understood as a literal appeal to a kind of mental entity, one that can be defined in terms of having *file structure*, and these entities can be made the subject of substantive explanatory claims. On the latter score, mental files are to be understood in the first instance as (part of) the cognitive basis of the paradigmatic inferential dispositions; they can then be made the target of substantive commitments that go towards providing an indication of sameness of reference. Although the first part is an instance of theoretical stipulation, the substantive commitments are made in the course of the second part (§5.2).

Treatments of mental files in the literature divide roughly in to those that make use of the *basic model* of mental files, one that only requires them to have file structure, and more elaborate models that impute to them more substantive properties. Examples of the first kind of explanation can be found in Lawlor and Schroeter. I will argue they are committed to psychologically unrealistic claims concerning the cognitive capacities of thinkers capable of referentially coordinated thoughts (§5.3). Recanati's more robust *indexical model* of mental files holds mental files to be governed by functional restrictions on information that individuate types of mental file analogous to types of indexical expression in public languages (§5.4). Although Recanati makes clear that he thinks this model is adequate to explain facts about rational inference, it is not clear what he has in mind. The most plausible application of the indexical model is that the functional restrictions that individuate file types supply an indication that the information comes from the same source (§5.5). The problem with the indexical model is that it provides, at best, an insufficiently general explanation of referential coordination, and perhaps none at all. The core cases of referentially coordinated thought have no plausible restrictions on which information can be associated with the relevant file, and there are arguably *no* mental files with such a restriction (§5.6).

5.2 *Explanation by Mental File*

In one form or another, the appeal to files has been made in a number of areas. In a useful summary, Salis (2013) locates the first application of mental files to philosophical theorising to Grice (1969), who discusses 'dossiers' in the context of making claims about empty names and the referential use of descriptions. Strawson (1974), building on Lockwood (1971), used files to frame claims about identity statements. Around the same time, Donnellan (1970, 1974) used files to discuss the semantics of proper names. Later work by Evans (1982), Perry (1980, 2001), Devitt (1989), and Forbes (1989, 1990) applied the idea in order to deal with related problems. Talk of mental files is an enduring feature of recent philosophical literature on thought and reference (cf. Jeshion 2009; Dickie 2010; Friend 2011). Files also figure in work in linguistics, notably in the work of Karttunen (1976) and Heim (1983), and in cognitive psychology, notably in the work of Kahneman, Treisman and Gibbs (1992).

Although these are all topics related to the problem of referential coordination, rather than go over this literature in detail, it is more useful to come at the question of how

explanation by mental files works, or could work, from a slight remove. I will do this by considering two general worries about explanation by mental files that can be raised even in the absence of any detailed statement. The first worry is that talk of mental files in any venue can only be *metaphorical*. The second worry is that, if they are to be used to give an explanation of referential coordination in particular, they seem to be constructions out of, or merely reflections of, what is supposed to be explained, and so talk of mental files is explanatorily *vacuous*. Both worries loom large over any appeal to mental files in the course of giving an explanation. Since the positive view that I will present in the next chapter makes use of mental file theory, it is worth spending some time on this. Saying something to ease these worries will help shed light on what mental files, and so what an explanation of referential coordination by mental file, will involve.

5.2.1 File structure

Salis (op.cit.: ii) notes, though without much emphasis, that talk about files in the mind is often taken as metaphorical. Recanati, one of the more prominent mental file theorists, says something to this effect (2012: vii), for instance. Even considered entirely in the abstract, there is obviously the potential for theorists to differ in how they understand talk of files to be useful for thinking about referential coordination, or any other feature of thought or language. Files as metaphorical entities could provide a way to model certain features of thought. Or else the existence of mental files might be accepted as a kind of fiction, though doubts about their genuine existence remain. Perhaps we ought to think that this is the right spirit, perhaps the most charitable spirit, to take talk of mental files, since the hedge of deeming them metaphorical protects talk of mental files from being rejected on the grounds of unfounded speculation about psychology. This immediately presents a worry about the suggestion that a genuine explanation can be given in terms of mental files. While metaphors are useful for explanation insofar as they serve as guides for what can be said if we want to say what is going on, if we want to know what is *really* going on, explanations in literal terms are needed. So if mental files are metaphors, then they may be seen as suspect as a means of providing a genuine explanation.

The idea that talk of mental files is best taken metaphorically is not clearly right, however. Talk of mental files can be understood as making literal appeal to a category of actual mental particulars with a certain structure that make an actual difference to the workings of minds of which they are a feature. For this literal appeal to be possible, a way of saying

what a mental file is that does not depend on simply saying that they are like normal files with the exception of being mental is needed. The key feature of the occupants of office drawers and filing cabinets is that they are particulars in which one can put various bits of information and from which one can take information, and yet they persist even so. This key feature as a guide can be used to spell out the notion of a mental file as follows:

Something is a mental file if it is (i) a mental particular (ii) that is associated with bits of information (iii) in a way that organises that information, (iv) has persistence conditions that do not depend on those associations (i.e. survives gain and loss of associations with bits of information), and (v) can enter into causal relations with other mental particulars, states, and events in virtue of the information with which it is associated.

Meeting conditions (ii-iv) is sufficient for having what we can call *file structure*. This is what mental files have in common with their non-mental counterparts. Conditions (i) and (v) say what is distinctive about mental files, that they are things with causal roles in the cognitive system, and that these roles are a consequence of their having file structure. What I refer to as mental file theory is the idea that there are mental entities with file structure.

File structure can therefore be defined, not just by saying that it is whatever mental files have in common with their non-mental counterparts, but by defining a general way for things to be that is common across otherwise different sorts of things. Perhaps this is a *kind* of metaphor, but it presents no kind of hedge as to the existence of files. Talk of mental files can be intended just as literally as talk of non-mental files. Consider the sorts of files that can be found on a computer: these are not the same as files that exist in the three dimensional world, but they are files all the same. They contain information, can exist even when that information is changed, and serve to organise that information, and so have file structure. Computer files are not metaphorical; they really exist. Talk of mental files can be intended as talk about really existing things in the same way.

The most interesting way of talking about mental files in the context of an explanatory project is this literal way. Explanations in terms of mental files should be taken that way, then, at least in the first instance. This remains the case even though the description of mental files is incomplete in some ways; there are many different ways of thinking about

information, association with information, and what the structure of the information so associated might be like. There are also different ways of understanding the causal relations that they enter into, perhaps, computational, associational, or both, or neither. File structure is realisable, as it might be put, in multiple ways. But it remains something that can be literally instantiated in the mind, however realised.

5.2.2 Two-part explanations

This brings up the second worry, that explanation by mental files is vacuous. The basis of the worry is the thought that mental files can only be defined in terms of whatever it is to be explained. There is some truth in this, though not in a way that makes explanation by mental file vacuous.

Fine gives forthright expression to the worry in the context of his discussion of coordination:

[I]t is hard to know what talk of mental files is meant to convey. Perhaps one thing it may reasonably be taken to convey is that certain items of information are stored together in a single “location,” while other items of information are not. Thus the information that Cicero is a Roman and that Cicero is an orator will be stored in the same location, while the information that Cicero is a Roman and that Tully is an orator (for someone who does not know that Cicero and Tully are the same) will not be.

But we may now ask: in virtue of what will information be stored in the same location or in a different location? After all, there is nothing intrinsic to the idea of co-location which requires that colocated items should be related in any particular way. And surely the answer to the question is that the location will be the same when the information represents its object as the same. (2007: 67-8)

Fine draws the conclusion that

mental files should be seen as *a device for keeping track* of when objects are coordinated (represented as-the-same) and, rather than understand

coordination in terms of mental files, we should understand the workings of mental files in terms of coordination. (ibid., emphasis added)

If this is right, then talk of mental files adds nothing to an explanation, since it presupposes what is to be explained.

A similar worry exercises Recanati in a related context:

Integrative behaviour on the part of the subject is a symptom of informational clustering, but it is not what constitutes, or accounts for, informational clustering. So the question: ‘What is it for two pieces of information to occur in the same file?’ cannot be answered simply by appealing to the subject’s integrative behaviour. On the mental file account, the subject’s integrative behaviour is explained by the hypothesised clustering of information. The further question, ‘what explains the clustering?’, still awaits an answer, and this is where circularity looms.... It is hard to answer that question without appealing to the fact that the subject takes the relevant information (that which goes into the file) to concern one and the same object. (2012: 95-6)

Recanati draws the conclusion that the worry that mental file explanations are circular

is essentially right and should not be resisted; but I also think it is a mistake to view the objection as undermining the mental file framework. One cannot, without circularity, account for internal co-reference (the fact that two pieces of information are taken to concern the same object) in terms of the occurrence of that information within a single file. If anything, it is the other way round: Two pieces of information go into the same file if they are taken to concern the same object. (2012: 101)

If this is right, then mental files presuppose the referential coordination of thoughts they are supposed to explain. Roughly the same idea is endorsed by Lawlor (2001: 80). Fine and Recanati’s ways of supporting the worry are slightly different, but both suggest that there is to a greater or lesser degree a problem with explanation by mental file being vacuous.

For Recanati, as a proponent of mental file theory, this seems like a less than ideal conclusion to reach. Be that as it may, we need not share it. How might talk of mental files bear on the explanation of referential coordination? It is possible to answer this in a way that shows the circularity worry to be based on a conflation.

The problem of referential coordination concerns the rationality of the paradigmatic inferential dispositions. This combines two different things to be explained. One is the *inferential disposition*, the fact that thoughts which are referentially coordinated are those with which the thinker is disposed to perform the paradigmatic inference. The other is the *rationality* of the dispositions, the fact that these dispositions do not just happen to a thinker, but instead there are reasons to have them. The inferential dispositions can be taken as one part of the causal properties of a cognitive system, explained by the sub-personal entities and processes that account for the make up of a person's mind. The rationality of the dispositions calls for explanation of a distinctive kind, one that says what reasons the thinker has. In the case of referential coordination, this is the indication of sameness of reference. Our focus is on the rational dimension, but the dispositional dimension need not be ignored.

Separating out the dispositional and the rational dimensions of referential coordination in this way provides the possibility for a particular form of explanation: one can introduce, by means of theoretical stipulation, some entity (and/or process) as part of the cognitive system to account for the dispositions, and then complete the explanation by providing substantive claims about those entities addressed to the way in which these dispositions are rational, thus giving the theoretical stipulation empirical content. Call explanations of this form *two-part explanations*. Two-part explanations based on theoretical stipulations has precedent in other areas; unobserved or 'hidden' entities with specified causal powers are postulated as what brings about a certain range of phenomena, and earn their keep through the substantive claims they can sustain and the gain in explanatory power thereby achieved.

Explanation by mental file, as I will understand it, is a two-part explanation. The first part is the introduction of mental entities with file structure by theoretical stipulation as the grounds of the disposition to engage in the paradigmatic inferences at issue. The inferential dispositions are causal consequences of a particular state obtaining in the thinker's cognitive system, of some bits of information being associated with a mental file (one might compare this with the introduction of structured mental representations to

explain causal facts about intentional mental states in the manner of the LOT hypothesis; I will return to this in §6.2.5). In the second part, these mental particulars are made the subject of substantive commitments so that some indication of sameness of reference is, at least in part, provided by the fact that the thoughts involve the same mental file. The content of these commitments gives the particular character of specific treatments of mental file theory. The desiderata on a good explanation of referential coordination apply here.

This two-part form of explanation is not the only way, nor the usual way in the literature, for mental file theory to be applied, though Recanati (2012: 96) says something along these lines (see also Schroeter 2012). Given this, it is puzzling that Recanati reaches the less than hopeful conclusion that he does. His reason, stated in the above quoted passage, is that mental file theory needs to say what makes it the case that two bits of information are associated with the same file, and that one needs to appeal to the fact that the subject takes the relevant information to be about the same thing to do this. We can now see that this runs together different senses of ‘makes it the case’ that should be kept apart. That a thinker has a certain set of inferential dispositions involving two thoughts *implies* that they are referentially coordinated and so, according to the stipulation, involve information associated with the same mental file. But it is not what *brings it about* that they are referentially coordinated. According to the first part of the explanation, being associated with the same file has the causal consequence that the inferential dispositions obtain. So this will be something concerning the implementation of the file structure, which can be set aside for the purposes of characterising the functions of mental files. Recanati’s conclusion only goes through if these two senses coincide, which they need not.

For the same reason, Fine’s objection that mental files merely track when thoughts are referentially coordinated is mistaken. What *makes* two bits of information ‘co-located’, or associated with the same file, is not that they are referentially coordinated, even though that is what, on the two-part explanation, *implies* that they are. Being co-located is a part of the explanation of the dispositional dimension of referential coordination. This makes use of a theoretical stipulation, but that does not make explanation by mental file vacuous. The substantive claims about mental files that provide the explanation of the rational aspect will be empirical claims about features of a category of mental particular.

Is explanation by mental file really of the two-part variety? One might think that the second part is redundant, or at least, it is obvious what it should look like, and so the distinction is only useful to show that the worry about vacuousness is mistaken. If mental files are associated with information, it might be said, then they have a topic or subject matter, and if they have a topic, then we have enough to say that mental files function so as to gather information on the same thing, and this function provides a defeasible indication of sameness of reference.

It is not at all obvious that mental files have the envisaged function in the relevant sense. Talk of 'function' is ambiguous between factive and normative senses: there is what something *in fact* does, and what it is *meant* to do. Only a normative function could provide the needed indication. Two mental files might in fact be associated with information about the same thing, but it is not part of what they are meant to do that that is the case. What is needed is an indication of sameness of reference, which would be an indication that the non-normative function is being carried out. Nothing about the idea of file structure as such bears on this.

The main point is that we need an answer to the question of *how* they get to have that function, and how it provides an indication of sameness of reference. Providing such an answer would just be to provide the second part of the two-part explanation. Making this simply a matter of stipulation would be to decline the task of offering a substantive claim about mental files to give mental file theory its empirical content. And since different accounts of how mental files explain referential coordination provide different comparisons with competing theoretical approaches, whether a positive evaluation of mental file theory can be returned will depend on whether any of these accounts supply an explanation without recourse to this sort of stipulation. Mental file theorists, therefore, would be advised to provide the substantive claims as part of showing why such an approach is explanatorily better than the alternatives, and go with the two-part form of explanation.

5.2.3 Summary

Mental file theory is the literal appeal to a class of mental particulars with file structure. Explanation by mental file can be given as a two-part explanation, the first part being the introduction of mental files through theoretical stipulation, the second being substantive claims about mental files that bear on the rationality of the paradigmatic inferential

dispositions. Neither point is compulsory, but they have clear advantages in responding to the two general worries about explanation by mental file, and so I will proceed on that understanding.

5.3 *Basic Model Explanations*

A mental file-theoretic explanation of referential coordination of the two-part variety needs to provide a claim about what feature or features of mental files provides an indication of sameness of reference. We can call the *basic model* of mental files the idea that mental particulars with file structure have no further explanatorily relevant properties considered in themselves. *Basic model explanations* employ only the basic model of mental files, and so need to bring in some further entities, processes, or properties to explain referential coordination. More robust models make comparatively stronger commitments, to the effect that mental files have functional properties in addition to the possession of file structure. Basic model explanations are at a significant advantage so far as discharging the demand on mental file explanations is concerned, being comparatively conservative in their commitments. However, as I will argue in this section, the weaker set of commitments about mental files seems to require a stronger set of commitments elsewhere that make the explanations psychologically unrealistic.

5.3.1 File management and metasemantics

How might a basic model explanation work? Some suggestions can be drawn from two treatments of mental files in the literature. Although neither treatment is directly addressed to referential coordination, they are both developed in response to similar problems, and so I propose to apply them to the task of providing an indication of sameness of reference.

According to Lawlor (2001), thinkers are in possession of mental files just in case they have *file management dispositions* (op. cit.: 79-91). By file management dispositions are meant dispositions to screen and prune information, i.e. to include only information that coheres with other bits of information and exclude information that does not, and to revise information on the basis of new information. File management dispositions (or perhaps their exercise) are then said to constitute a reliable attempt to *maintain an intentional relation*, i.e. sustain the information associated with the file as being about a

particular thing (op. cit.: 72-5). File management dispositions that constitute such an attempt therefore provide an indication of sameness of reference.

Schroeter (2012) argues that the fact that information associated with a mental file is treated as being about the same thing means that they are the target for the application of what she calls a *metasemantic theory* (op. cit.: 190). A metasemantic theory in her sense is a set of propositions that says that thoughts that meet certain conditions should be *interpreted* in a certain way, as being about one thing or another, or nothing. Part of this theory is the axiom is that if a thinker is disposed to treat her thoughts as being about the same thing, then all things being equal, they are about the same thing. The ‘all things being equal’ clause signals the fact that this is one of several axioms which must be taken as a whole. These will presumably also make reference to other facts about the thinker’s dispositions, although Schroeter does not provide a detailed treatment of potential defeaters for the interpretation. The explanation from this treatment is that the indication of sameness of reference comes from the fact that the dispositions and the metasemantic axioms combine to make it the case that the information is in fact about the same thing.

Schroeter explicitly emphasises that she is attempting to provide what she calls a *connectedness model* (op. cit.: 190-5), which means that it provides an explanation of the rationality of a thinker’s inferential dispositions in terms of relations between thoughts, rather than simply resemblances between thoughts that depend on properties that can be specified independently. In a similar way, Lawlor’s idea of file management dispositions is a relational fact, having to do with what a thinker does with her files, rather than in properties of the files considered independently of this. Both explanations therefore depend on some cognitive capacity thinkers have with respect to their mental files such that information’s being associated with the same mental file is an indication of sameness of reference.

It is important to emphasise that it is not merely the obtaining of the dispositions, but a further fact about a thinker’s cognitive capacities that provides the indication of sameness of reference on these basic model explanations. On Lawlor’s approach, the exercise of file management dispositions constitutes a reliable attempt to make the information bear on the same thing. This is important, since it cannot be that the file-management dispositions themselves do the work. The thinker must be sensitive to the dispositions as constituting the attempt, as Lawlor recognises (2001: 91-95), otherwise it is not clear

why file management dispositions would bear on the rationality of the paradigmatic inferential dispositions. As Lawlor has it, “sensitivity to one's files is a capacity for detection of the presence of one's own characteristic screening and pruning dispositions.” (op. cit.: 95) Similarly, if the metasemantic theory does not figure in the characterisation of any cognitive capacity that the thinker has, it is unclear how their application to the information that the thinker is disposed to integrate should make any difference to the rationality or otherwise of that disposition.

5.3.2 Psychological realism

Both basic model explanations rely on a distinctive sort of cognitive capacity, a capacity that provides for sensitivity to the fact that information associated with a mental file is about the same thing. There is no independent reason to think that all thinkers who are capable of referentially coordinated thoughts have these capacities, and there are good grounds for thinking that they do not. So the explanations fail to meet the psychological realism desideratum.

Sensitivity to one's file management dispositions or metasemantic axioms should be evinced by the thinker if the explanations are going to work. This awareness need not be in the form of something that the thinker could *express*, and so it is not to the point to say that thinkers are not (normally) able to talk about their file management dispositions or metasemantic theory. It just means that there needs to be something in the dispositions or activities of the thinker that provides evidence for this sensitivity, and it is not clear that there is any such evidence. There seem to be a couple of options for what counts as evincing this sensitivity, neither of which are good ones. If it is argued that the rationality of the paradigmatic inferential dispositions is itself evidence for the role that cognitive policies or metasemantic theory play in a thinker's dispositions, then the account becomes circular. If it is denied that there needs to be any such evidence, then the explanations become open to the charge that they fail to provide an adequately personal-level explanation. If the sensitivity is operative at all, it would not be operative at the personal-level, and so it would become unclear how it would provide the thinker with reasons to treat her thoughts as being about the same thing.

So, as has been the case with many of the problems raised against the explanations that have been examined, a charge of failing to meet a desideratum or requirement can seemingly only be avoided at the cost of failing to meet another.

Since neither the file management dispositions treatment nor the metasemantic axioms treatment is addressed to referential coordination as such, this should not be taken as internal criticisms of either Lawlor's or Schroeter's discussions. The basic model explanations in terms of file management dispositions and metasemantic axioms are extensions of their treatments to a different explanatory concern. One might be tempted, in light of this, to object that the charge of psychological unreality is nullified by the accounts' explaining other phenomena. Although the explanatory concerns are different, they are related, and this objection fails as a result.

Schroeter focuses on coreference *de jure*, which is stronger than referential coordination, but weakens it to make a condition that can be taken as coincident with referential coordination; the idea is that the thinker has reason to take her thoughts to be *de jure* coreferential, which is just what referential coordination is. Lawlor sometimes puts her explanatory aim in terms of showing what warrant a thinker has for treating her thoughts as having the *same sense*, though it is not clear if she means this to pick out the thicker theoretical notion of sense discussed above. If so, then it is redundant. If not, then it is just the same as what that thicker notion is meant to explain. Moreover, it is clear from Lawlor's remarks that she takes this kind of warrant to be something that is only had if it is rational to treat the thoughts as referring to the same (§6.4.2 below provides related discussion on these points).

5.3.3 Summary

On the present showing, what basic model explanations gain by using a simpler model of mental files, they pay for with a less realistic view of thinkers' cognitive capacities. One could extrapolate from the problem raised for the Lawlor- and Schroeter-inspired accounts that any way of spelling out the relationship between mental files and the cognitive system in such a way as to give a basic model explanation will face the same problem. It is one thing to say that there are such things as mental files, but another to say that we have dealings with them that are exercises of certain cognitive capacities, and that these exercises are what indicates sameness of reference in core cases. The cognitive capacities in question are not something we have reason to think we possess independently of the need to provide a basic model explanation.

Ultimately, this line of reasoning rests on a conception of how one might employ the basic model of mental files that is too narrow, and an idea about the relevant cognitive capacities that is too strong. Something similar to, though much less demanding than, Lawlor's view can provide an adequate explanation of referential coordination that avoids the charge of being psychological unrealistic. And Schroeter's idea of a connectedness model is essentially correct in its idea that it is not resemblances between thoughts but relations between them that depend on what is done with them that matters (a idea also to be found in Millikan 2000, see §4.7). For the time being, however, I will treat the worry about basic model explanations as motivating more elaborate models of mental files.

5.4 *The Indexical Model*

Recanati's (2012) detailed treatment of mental files grasps the nettle and commits to a far more robust take on mental files than the basic model. The *indexical model* construes mental files as vehicles of thought akin to indexical expressions in public language. The key feature of the model is the idea of functional restrictions on information associated with files. This is the central part of a somewhat complex overall view of mental files that the indexical model encompasses. I will set out this view in as much detail as needed to see how it might supply an explanation of referential coordination, and then argue that the most plausible of these is nevertheless not a good explanation of referential coordination. The problem with the explanation provided by the indexical model is that, while it relies on mental files having functional restrictions on information, the core cases of referential coordination need to be understood as featuring mental files without functional restrictions, and so the explanation does not apply.

5.4.1 Functional restrictions and indexicality

Mental files are understood by Recanati in a way consonant with the definition of basic mental files given in the previous section, as things that store information (2012: 33) and survive changes in informational association (op. cit.: 35-38). However, he also claims that they are vehicles of thought, and that their being such is a consequence of their having certain functional properties.

Recanati's frames his presentation of the indexical model with the distinction between *general* or merely descriptive thought about things, and *singular* thought, thought that is

not about something because of its satisfaction of a description. The viability of this distinction is a matter of some controversy (the papers in Jeshion 2010 provide a number of recent interventions in the various debates concerning this distinction), though we need not go into this for present purposes. Following Bach (1987: 12), the distinction is understood as that between *satisfactional* and *relational* denotation; X denotes Y satisfactionally iff. there is some condition that X denotes and Y satisfies that condition. X denotes Y relationally iff. there is a relation that obtains between X and Y, or alternatively, between the producer/thinker/speaker of X and Y.

The kinds of relations in which one can stand to something in order to have a singular thought about it are acquaintance relations, or what Recanati (2012: 20, 34-39) calls *epistemically rewarding relations* ('ER relations' for short). These are relations between a thinker and an individual such that the thinker can get information about that individual. That singular thought depends on acquaintance is a view, often traced to Russell (1911), which is widely accepted, though rejected by some (see Jeshion 2009; Hawthorne & Manley 2012 for recent discussion in this vein). Recanati's view is comparatively liberal with respect to what counts as an acquaintance relation when compared with Russell's, even to the extent of allowing ER relations that do not rely on any particular form of getting information (on which much more below). Because mental files are vehicles of thought that are based on ER relations, and because ER relations determine the referent of the file, there is some blurring of the content/vehicle distinction. Some of Recanati's critics (e.g. Pagin 2013: 136-7) make heavy weather of this, though it is not obvious that an un-blurred version of this distinction is something that we should insist upon (cf. Recanati 2013: 213-8). But it is worth noting that the indexical model is intended as a theory of the vehicles of thought, and so is heavyweight.

On the indexical model, mental files are said to be *based on* ER relations. Being based on an ER relation explains why and how mental files are vehicles of (singular) thought. But it is not immediately clear what it means. Recanati talks about this aspect of ER relations in two ways: first, the information associated with a file *should have* been gained through the relevant ER relation (2012: 61); second, the mental file ought to exist *only so long as* the relevant relation obtains (ibid.).

According to the first claim, mental files are governed by what we can call *acquaintance norms*. According to the second claim, files are also governed by *existence norms*. How are these claims related? According to Recanati,

Since [/if] the function of a (type of) file is to exploit a given (type of) ER relation, [then] a token of that type should come into existence only if the subject stands in the appropriate contextual relation to some entity, a relation in virtue of which it will be possible for him or her to gain information from it. Unless there is an information channel of the appropriate type which the file can exploit, there should be no tokening of that type of file. (op. cit.: 52)

The idea is that mental files are governed by acquaintance norms because files are created (variously ‘opened’, or ‘tokened’) in order to exploit particular ER relations. So acquaintance norms are supposed to apply because existence norms apply.

The first way of taking the idea that mental files are ‘based on’ ER relations, that they are governed by acquaintance norms, is the more important for understanding the indexical model, and why it is *indexicality* that is the important property. Acquaintance norms specify a kind of ER relation, the norm being that only information got via that ER relation can be associated with the mental file. Being governed by an acquaintance norm is what I will call a *functional restriction* on the information associated with mental files. Indexical expressions, such as ‘I’, ‘here’, and ‘now’, are expressions with a *standing meaning* that does not vary across occasions of use, and a *token reference* that is fixed by particular contexts of particular use. Functional restrictions supply the analogue of standing meanings. The things that stand at the other end of the ER relations are the referential contents, the analogues of token reference.

Recanati sometimes (e.g. op. cit.: 60) makes this claim by saying that files are *individuated* by ER relations. The indexical model appeals to the distinction between file types and token files. Since types and tokens are different sorts of entity, the individuation by ER relations can mean different things: either that ER relations distinguish types of mental file, or that what type a token mental file belongs to depends on the ER relation it is based on. The idea that ER relations ‘individuate’ token files is not wholly accurate. Standing meanings do not individuate token indexical expressions (two tokens of an indexical expression, for instance ‘I’ and ‘I’, might have the same standing meaning, but are different) but are individuated by their standing meaning and also whatever makes them individual expressions more generally. Similarly, token files will be individuated by whatever makes them individual mental entities. This point

makes clear that indexical mental files are similar to but not the same as the propositional constituents posited by sense theory (Recanati's discussion sometimes plays down this point, cf. 2012: 27-41, 244-6). They are mental entities, and so have a different sort of individuation condition. In light of the problems that were raised for sense theory, this is important. Token indexical mental files might be seen as 'quasi-senses'; they share some properties with senses, in particular that of determining referential content, but are different most obviously in being mental particulars.

5.4.2 The hierarchy

Just as token files are in this way based on ER relations, different types of indexical mental file are based on different kinds of ER relations. There are three broad classifications of ER relations germane to this question: *basic* ER relations, *composite* ER relations, and the *higher-order* ER relation. The differences between and details of these and the functional restrictions they impose through acquaintance norms are important for evaluating the indexical model, and so it is worth going through what Recanati says about them.

Recanati discusses a number of examples of basic ER relations. *Perceptual demonstrative* files are those that exploit present perceptual ER relations, information links to objects made available through various sense modalities, normally though not necessarily those that are in one's present surroundings. Different kinds of perceptual ER relations can be distinguished according to different modalities. *Memory demonstrative* files (op. cit.: 62) are those that exploit past perceptual ER relations, those that do not hold at the time the file exists storing rather than being updated with information gained through perceptual contact. *Recognitional* files are those that are based on an ER relation Recanati terms *familiarity*, which depend on an ability to recognise an object due to 'multiple exposure' to it. Familiarity is context-dependent in the same way as past and present demonstratives, because "[i]n a different environment the very same recognitional device in place in the subject would have had the function of detecting another object than what it actually has the function of detecting in the actual environment." (op. cit.: 59). Whether this makes it genuinely context-dependent depends on it being the case that one and the same recognitional ability could pick out different things in different contexts.

Although each is defined to some extent with reference to the others, the basic ER relations and the files that are based on them cannot be reduced to any more basic ER relations or type of files. *Composite* ER relations, by contrast, are the result of a file gaining information through *combining* or *compounding* (op. cit.: 71) basic ER relations. Unsurprisingly, the combinations are many and various: one can have a *perceptual recognitional* or *memory recognitional*, or more generally *demonstrative recognitional*, as well as *memory-perceptual* ER relations. These count as non-basic recognitional relations because they are not based on the basic familiarity relation, but a combination of basic ER relations that has a similar effect to the basic familiarity relation.

A classification of ER relations that cuts across the basic/composite distinction is the distinction between *stable* and *non-stable* files. An ER relation is stable when the individual picked out does not necessarily change given a change in a thinker's context over time, a change in what one can perceive or where one is in the world. Basic recognitional files are stable, whereas demonstrative-recognitional files are not, since demonstrative files are not. A particular species of stable file is what Recanati terms the 'self' file, which is based on ways of getting information that one has only in relation to oneself, i.e. proprioception, nociception, etc. (op. cit.: 68-75).

The third kind of ER relation, the *higher-order* ER relation, is the relation that obtains whenever any of the first-order, i.e. basic or composite, ER relations obtain. Recanati calls files based on higher-order relations *encyclopedia entries*. Encyclopedia entries are thought of as equivalent to or doing the work in thought of proper names (Recanati discusses them in this way in his 1993: 181-187). What is the point of such files? One answer is that there are many things we can think about for which the ability to think about them does not depend on any perception, memory, familiarity, nor any combination of these. Nor does this ability come with any particular restrictions on ways that information about them be gathered. Rather, such thoughts exploit whatever way of getting information about the object in question are available in an 'opportunistic' fashion (cf. Recanati 2012: 73).

Proper names are naturally taken to be quite different from indexical expressions; similarly, mental files for opportunistic thoughts appear not to have the features of indexicals that indexical mental files are supposed to have, there being no obvious analog to standing meaning as distinct from token meaning; what is the type of which they would be tokens? However, Recanati writes,

I think we can fit encyclopedia entries into the indexical model. The only thing we need is to make room for a new sort of relation to the referent: a higher-order relation which holds whenever we stand in some ER relations to it. In other words, rather than say that some files are not based on relations to the referent, I draw a distinction between specific ER relations and the higher-order ER relation on which encyclopedia entries are based, namely

$$\lambda x \lambda y [(\exists R)(Rx, y)]$$

where 'R' ranges over ER relations. A subject (or a mental file in the subject's mind) x stands in that relation to an object y just in case there is/are some first-order ER relation(s) in which x stands to y. A file based on the higher-order relation hosts any information derived in virtue of that relation, that is, ultimately, any information derived in virtue of any of the first-order ER relations.... Encyclopedia entries come at the top of a hierarchy of files. (op. cit.: 74-5)

The indexical model therefore provides a type of mental file for cases where the way someone has of thinking about something is not dependent on particular ER relations she stands in to them.

5.4.3 Operations

The hierarchy of which encyclopedia entries form the top is articulated in terms of *operations* on files. One operation in particular is relevant to the explanation of referential coordination, that of *conversion*, "the process through which information stored in a file is transferred into a successor file when the ER relation which sustains the initial file comes to an end." (op. cit.: 81) Conversion comes in two sorts: it can involve change from one basic ER relation to another, or change from a basic to a composite ER relation. Recanati terms the second sort *incremental conversion* (op. cit.: 87). In addition to this, *expansion* is incremental conversion that operates on what Recanati calls *proto-files*: these are mental files that can only contain information gained through one ER relation. Incremental conversion turns these into *conceptual* files, which are those that can contain information gained in other ways, and so can be employed in thought more

generally. Recanati's remarks on how encyclopedia entries relate to proto- and conceptual files (i.e. first-order files) are sketchy, though he suggests that the creation of it might also be understood as a kind of incremental conversion, albeit one that depends on a thinker's linguistic capacities in some way (op. cit.: 102-3).

A less speculative suggestion, which is nevertheless consistent with the involvement of language, is that, along with incremental conversion, there is also *detachment*. Detachment is a species of conversion of a conceptual file based on a composite relation to an encyclopaedia entry by dropping all functional restrictions on the information. There will be the requirement that all the information comes from the same thing, but this will not impose any specific functional restrictions, like the exclusion of information gained through certain sensory modalities. Incremental conversion is a species of conversion operation that, unlike detachment, *adds* to the information about the information associated with a file by chaining additional ER relations. Detachment, by contrast, *takes away* information, by dropping the functional restrictions. Whether the creation of encyclopaedia entries is a matter of linguistically-mediated incremental conversion or not, what is important is that mental files based on the higher-order ER relation are not based on any particular first-order ER relation. This may seem like a minor detail, but it will play an important role in later discussion.

Recanati's treatment of the indexical model involves other operations that are more properly understood as incremental conversion, rather than detachment. These are *linking*, through which information contained in two files is shared, and *merging*, through which two files are fused to create one file. Linking and merging capture the way mental files relate to judgements of identity. Thinking that one thing is the same as another can be understood as relating two files in these ways, either by sharing information across them, or fusing the files themselves. These relations describe the dynamics of a thinker's system of mental files on the indexical model, though the general idea of operations on files has application outside of that model (see §7.3 for further discussion).

5.4.4 Explanation by indexical file

What explanation of referential coordination might we get from this? Recanati touches on related problems under the heading of 'integrating information', and offers this:

Files are a matter of information clustering. Clustering takes place when all the information derives from the same source, through the same ER relation, and when it takes place, it licenses the integration and inferential exploitation of the information in question. The role of the file is precisely to treat all the information as if it concerned one and the same object, from which it derives. But integration and exploitation of information is blocked if the relevant information is distributed in distinct files, for then, there is no presumption that all the information derives from the same object. So, even if I know that Cicero is bald, and that Tully is well-read, I cannot conclude that some bald man is well-read, despite the fact that Cicero is Tully: the information 'is bald' is in the Cicero file, while the information 'is well-read' is in the Tully file. Informational integration and inferential exploitation of information only takes place within files, on this picture. (2012: 42-3)

Although this passage might be read as presenting clustering as something that follows directly from the notion of a mental file ("files are a matter of information clustering"), clustering as characterised is not something that arises out of the notion of a *basic* file introduced above. Talk of clustering as happening when information derives from the same source risks conflating two things that need to be distinguished: *that* information is associated with a file, and *how it is* that information comes to be associated with a file. The idea of clustering therefore must involve the substantive claim that information is supposed to be associated with mental files because it was derived from the same source through the ER relation specified by the acquaintance norm on the file. This will bear on the fact that it is information about the same thing, and presumably this plays a role in 'licensing' the integration. But how exactly?

A natural suggestion is that it means that mental files themselves have referential contents on the indexical model. They are supposed to be vehicles of singular thought. If they are vehicles of singular thought, then sameness of mental file implies that either there is sameness of reference, or no reference at all. A problem with this is that it seems to require an identity judgement, albeit in part regarding the mental file involved in the thoughts, and what is wanted is an account of how identity can be represented without such judgements, since it would otherwise misdescribe the distinctive sort of inferences that are made available by referentially coordinated thought. The deeper problem is that it is psychologically unrealistic, requiring as it does the thinker to form

beliefs about the representational aspects of their cognitive system as such. This meta-representational capacity is one that we have no reason to attribute to a thinker independently of the proposed explanation. So the natural suggestion ought to be resisted on closer inspection.

An alternative explanation, and one that fits in more neatly with the overall thrust of the indexical model, is one that appeals to the acquaintance norms, not as what confers referential content on mental files, but as placing constraints on the information associated with mental files. These constraints provide information about the information, along the following lines: if there is a functional restriction on the information associated with a mental file, then there is an indication that the information was associated with the mental file as a result of being derived from the same ER relation; if the information was associated with the mental file as a result of being derived from the same ER relation, then the information comes from the same thing, since token ER relations are individuated in part by their relata; so if there is a functional restriction on the information associated with a mental file, there is an indication of sameness of reference. This explanation makes use of the fact that the ER relations on which files are based serve to identify objects. The indication of sameness of reference that the functional restriction provides is defeasible, but it need not be otherwise in order to be adequate to explaining referential coordination.

It is hard to say with absolute confidence that this a fair representation of Recanati's thought, especially since he spends practically no time spelling out how it is he thinks clustering licenses the integration of information. But I submit that the reconstruction in terms of acquaintance norms presents the most plausible way of joining up the indexical model with the idea that clustering has this licensing effect.

Critical responses to Recanati's indexical model have largely focused on concerns with how well it can be motivated, particularly with regards to comparisons with (what are effectively) basic model explanations (cf. Hall 2013; Papineau 2013; Ball 2015; Ninan 2015), and on whether the indexical model is coherent, particularly with regards to whether the indexical typing of mental files coherently maps onto the rationality of inferences across different ways of getting information or across information gained at different times (cf. Goodsell 2013; Onofri 2015). Recanati (2013) has responded to both sorts of worry, and it is not my intention to press either.

It is, admittedly, sometimes unclear exactly what Recanati thinks motivates the indexical model. One can grant as highly plausible the idea that files *stand in* ER relations, but reject the move from that to their being *based on* ER relations and so governed by acquaintance norms. But we can grant that the indexical model is well motivated so long as it provides a good explanation of referential coordination. On this score, it seems to avoid problems that beset other explanations. That there are ER relations, that there are different sorts of ER relations, and that there are functional restrictions of these different sorts are all plausible, and so no untoward demands are placed on thinkers in order to have referentially coordinated thoughts. What provides the indication of sameness of reference on the indexical model, the functional restrictions, are not things that are to be analysed, introduced, or explained in terms of referentially coordinated thoughts, so the non-circularity desideratum also looks to be met.

5.4.5 Summary

On the indexical model, mental files are governed by acquaintance norms that specify ER relations between the file and an individual. These functional restrictions are of different kinds, giving rise to different types of file. Files of different types can be related through various operations. The functional restrictions on files can explain referential coordination because the ER relations specified by the acquaintance norms obtain between only one individual and a file. This explanation avoids many of the problems with competing explanations.

5.5 *The Generality Problem*

Despite its merits, there is room for doubt about whether the indexical model explanation is sufficiently general. One of the desiderata on an explanation is that it should extend to the core cases. The indexical model, as the name suggests, is directed at peripheral cases, thoughts about ourselves, our present location and present time, which Recanati discusses at length. He also discusses cases where a thinker is getting information about objects with which she is in perceptual contact, for which the term ‘demonstrative’ is used (see the discussion of basic ER relations in §5.4.2). These cases are apt to be put under the heading of special ways of thinking. But the core cases are not like this. They fit into the indexical model as cases that involve encyclopedia entries, those at the top of the indexical model’s hierarchy. The problem is that the resources needed for indexical explanation are absent from this sort of file. A more radical worry,

which I shall sketch but not develop further, is that *all* cases where referential coordination is possible would need to involve encyclopedia entries. Either way, the indexical explanation faces a problem with generality.

5.5.1 Absence of indexical explanations from encyclopedia entries

Before setting out the problem, a bit of terminology. A mental file featuring functional restrictions on the information associated with it, as described by the indexical model, is an *indexical file*; an indexical file *has a first-order functional restriction* if and only if it is governed by an acquaintance norm that specifies a first-order ER relation; a *case* consists of a thinker with the paradigmatic inferential disposition involving two thoughts, and who is rational in doing so; an explanation of referential coordination that relies on the functional restriction on information associated with a mental file is an *indexical explanation*. An indexical explanation of a case involves the assumption that the thinker has an indexical file with which is associated information that is to be exploited in the paradigmatic inference, and for the sake of argument, we can grant that this assumption is otherwise free of problems.

The claim that causes problems for indexical explanation is:

Absence: There is no indexical explanation in cases where the indexical file is based on the higher-order ER relation.

Absence is supported by two principles concerning indexical explanations:

Dependence: There is an indexical explanation of a case only if there is a mental file that has a first-order functional restriction in that case.

Exclusion: If the mental file in a case is based on the higher-order ER relation, then that file does not have have a first-order functional restriction.

The problem is that indexical explanations work through functional restrictions on informational association, but mental files based on the higher-order ER relation have no such functional restriction. The two claims that imply Absence can be motivated by

looking at the details of the indexical model explanation in cases of mental files based on ER relations of both the first- and higher-order.

Exclusion follows from the nature of encyclopaedia entries. Encyclopaedia entries are created by the kind of expansion operation that I earlier termed *detachment*. Detachment is the operation of dropping first-order functional restrictions (either through creating a successor file or through modifying an existing file; as per the discussion above, this difference can be ignored for the purposes of this objection). In light of the distinction between proto- and conceptual files, these first-order functional restrictions will involve composite ER relations. Because encyclopedia entries, indexical mental files based on the higher-order ER relation, are supposed to be able to include information gained opportunistically, from any source, they exclude any composite ER relations that could figure in a functional restriction on information associated with the file.

To see why Dependence obtains, consider the following abstract description of a case. Suppose F is a perceptual-recognitional file. Call the functional restriction on the file R. Because F is governed by R, information associated with F is gained through a composite ER relation involving both a perceptual ER relation (call it P) and a memory ER relation (call it M) to the same thing. Information associated with F can be integrated by means of the paradigmatic inferences because R is a restriction on the information associated with F such that the information is gained through P and M and information gained through P or M should be from the same thing. The first-order restriction F involves the disjunction of two restrictions involving basic ER relations, which together serve to indicate sameness of reference. Without the first-order restriction, there would be nothing that would supply an indication of sameness of reference as per the indexical explanation. So if we removed those restrictions, we would remove F. And if we remove F, then there is nothing to indicate that the information comes from the same thing. Thus indexical explanation depends on there being a first-order functional restriction, i.e. Dependence.

Absence presents a problem for indexical explanation because it means it fails the desideratum of sufficient generality. Cases where we have more than one way, perhaps an indefinite variety of ways, of getting information on something, and can integrate the information gained in those several ways, are amongst the core cases. Encyclopedia entries, the mental files are suited to such cases, are based on the higher-order ER

relation. The argument from Dependence and Exclusion to Absence shows that indexical explanation does not extend to those cases.

5.5.2 Responses

Is there some way for the proponent of indexical explanation to avoid the charge of insufficient generality? They might reject Dependence, though it is not clear how there could be an indexical explanation that did not depend on a first-order ER relation. Although Exclusion follows from the definition of encyclopedia entries, a defender of indexical explanation might wish to respond by arguing that, in some way, there is a first-order functional restriction even in the case of encyclopedia entries:

Against Exclusion 1: Since the higher-order ER relation was introduced by abstracting over first-order ER relations, it follows that when an indexical file is based on the higher-order ER relation, there is a first-order ER relation in place; if there is a first-order ER relation between an object and the indexical file, then there is a first-order functional restriction on the file. So Exclusion can be rejected.

Something like this is quite possibly what Recanati has in mind when discussing the higher-order ER relation, for all the little he says in terms of the nuts and bolts of licensing the integration of information.

The first part of the response can be granted. Exclusion does not say that, when a file is based on the higher-order relation, there is no first-order ER relation in place. That would be in direct conflict with what the higher-order ER relation is. What it does say is that there is no functional restriction that specifies a first-order ER relation. The problem is that, even if there is a first-order *ER relation* in place, it is not the case that there is a first-order *functional restriction*. This is a problem with the second part, the idea that if there is a first-order ER relation between an object and the indexical file, then there is a first-order functional restriction on the file. Since it is not sufficient for indexical explanation just that there is some first-order ER relation, this is needed for the response to work. But it is also false. If there is a first-order ER relation whenever the higher-order relation obtains, which there is, then for it to provide an indexical explanation, it must be one that figures in a functional restriction. That requires that the file is based on that ER relation. An encyclopaedia entry, by definition, is *not* based on a first-order ER relation.

The move from there being the higher-order ER relation to there being a first-order functional restriction is incorrect, and so Exclusion can stand.

The only sense in which there is a first-order functional restriction is the one which is in the file's *operational past*, prior to the creation of the encyclopaedia entry through the detachment operation. This does not provide any reason to reject Exclusion, however. If the higher-order relation obtains, and so there is a first-order ER relation in the file's operational past, that first-order ER relation is *screened out* by detachment, in the sense that it no longer supplies a functional restriction on the information associated with the file. This is so however the conversion operation, of which detachment is one instance, is supposed to work, either as change in a numerically identical file, or change across numerically non-identical files. Information can be (that is, it would be permitted by the higher-order acquaintance norm to be) associated with that file in a way that would have been ruled out by a first-order acquaintance norm, and so it does not even partially govern a file based on the higher-order ER relation. A mental file that was once based on a first-order ER relation, or a file that is the descendant of a file based on a first-order ER relation, is no longer based on a first-order ER relation after detachment. There can, therefore, be no way of getting an indexical explanation in virtue of a norm based on a first-order ER relation in a file's operational past.

The response that denies Exclusion might be pressed in a slightly different way:

Against Exclusion 2: Although it might not follow from Recanati's explication of the higher-order ER relation that encyclopedia entries based on it have a first-order functional restriction, we can change how the encyclopedia entries are to be understood so that it *does* follow. So Exclusion is false.

This response depends on making a change to how the higher-order ER relation is understood. But what change, exactly, and why make it?

I take it that Recanati's explication of the higher-order ER relation takes the form that it does for a reason, and so this response implies that that reason is non-conclusive, and anyway is shown to be misleading by the problem posed by Absence. Perhaps this is all the motivation one needs, though it is apt to seem highly ad hoc. In any case, it cannot be made to work. The change would have to make it the case that, when the higher-order

ER relation obtains, there is a first-order functional restriction. There are two problems with this. One is that this would be the case only if there were some particular individual that the higher-order ER relation picks out as the source of *all* the information in the file. Now, it is true that the higher-order ER relation obtains when there is an individual that is picked out by a first-order ER relation, and also that the higher-order relation *depends on* there being at least one first-order ER relation between a file and an individual. This does not entail that there is one individual from which all the information associated with the file comes. So an indexical file's being based on the higher-order relation provides no indication that the same individual is picked out by any of the ER relations in which the file stands.

The change would therefore have to be that encyclopedia entries are not based on the higher-order ER relation, but some other kind of ER relation adequate to the task. The obvious suggestion is that they are based on a kind of composite first-order ER relation, albeit an open-ended and highly complex one. Call these *compendious ER relations*. But, and this is the second problem, compendious ER relations do not sufficiently differ from the higher-order ER relation in order to show that Exclusion is false. Compendious ER relations are disjunctions of indefinitely many first-order ER relations. Compendious ER relations would, in that case, supply no difference with respect to what functional restrictions are in place. Effectively, there is no functional restriction supplied by a compendious ER relation, since one can always make an exception for any information potentially associated with the file. Indexical files based on a compendious ER relation and those based on the higher-order ER relation would not differ in any meaningful way from the perspective of whether an indication of sameness of reference is supplied. If indexical files based on compendious ER relations are functionally no different from encyclopedia entries, it still does not follow that encyclopedia entries allow for indexical explanations. So Exclusion can stand.

The only other strategy available to the proponent of the indexical model is to deny that Absence is really a problem after all:

Absence Is Not a Problem: Indexical explanation is limited to cases of mental files governed by first-order ER relations, as per Absence, but Absence is not a problem because (i) there are no encyclopaedia entries; (ii) encyclopaedia entries are not that common; (iii) whether or not they exist

or are common, we cannot rationally integrate information associated with encyclopaedia entries; (iv) other explanations are no better off.

These four different ways of denying that Absence is a problem differ greatly in their plausibility: (ii) and (iii) are both false, since, *ex hypothesi*, encyclopaedia entries must be common enough and it must also be possible to integrate information associated with them rationally, otherwise they would be members of a category of mental entity that lack the feature that motivated positing them; the fate of (iv) remains to be decided; the next two chapters will refute it; (i) perhaps provides a way of denying the problem without falling in to the problems with the second way of rejecting Absence just considered, though it needs elaboration. The thought might be that all we *really* need are files governed by (not compendious ER relations or the higher-order ER relation but) simply highly complex composite first-order ER relations. While Recanati mentions composite ER relations consisting of two basic ER relations, the indexical model itself is compatible with far more complex composite ER relations, and these need not be open ended. The indexical model can therefore be absolved from the charge of insufficient generality if this extension of composite first-order ER relations can be made to do all the work. Or so goes this reply.

There are two points to be made in response, one outright and one more speculative. The outright point is that the indexical model gains no advantage by relegating encyclopedia entries to being an optional extra. Encyclopedia entries are needed so that the indexical model is adequate for the task of describing (never mind explaining) rational thought and inference, and the ways we have with dealing in information, which includes the sort of opportunistic way of getting information about things that is, after all, why Recanati is moved to extend the model to include the higher-order ER relation in the first place. There is a way of *having* information that can only be adequately captured on the indexical model by positing the higher-order ER relation, the way that involves no functional restrictions on the information because it is not tied to any particular way of *getting* it. What matters is that we have it, not how. As Recanati says, this sort of file is needed because, in some cases, we can gather information opportunistically; any ER relation will do, and so the opportunistic gathering of information requires the dropping of functional restrictions on information.

In this connection, it is worth noting that Absence says nothing against the coherence of the basic idea behind indexical model. Nor does it say anything about the higher-order

ER relation as such. If there are ER relations, and it seems hard to deny that there would be if there are mental files, then the higher-order ER relation obtains just by definition. Moreover, the higher-order ER relation can be used to coherently extend the indexical model to the sorts of cases where what information we have about individuals is not gained from any specific source, or so we can suppose for the sake of argument. The point is that files based on the higher-order ER relation do not bring along with them acquaintance norms that can provide indexical explanations, and that is an obvious problem for the explanatory scope of the indexical model. Casting files based on the higher-order ER relation as optional extras does not avoid this problem, but exacerbates it.

The more speculative point is that there are reasons to think the operation of expansion is coextensive with the operation of detachment. Any file that is converted through expansion is thereby converted through detachment. In which case, it is not just adequacy in covering cases that makes encyclopedia entries necessary, but the nature of the dynamics of mental files in all the relevant cases. The reasons are these:

- (1) If an indexical file is expanded, then it can be involved in global inference.
- (2) If an indexical file can be involved in global inference, then it may include information that was not gained through any particular ER relation.
- (3) If a mental file may include information that was not gained through any particular ER relation, then it must be detached.

Indexical files are expanded so that they can be used in gaining information about how the world is and to guide our actions in light of that information. One of the ways of gaining information is global inference, by which I mean the kind of reasoning wherein *any* information a thinker has is potentially relevant, and potentially in need of update as a consequence of reaching conclusions. There is, notoriously, a problem of *how* this is achieved, and even a question of the extent to which it is unrestricted, but that it *may* be rationally undertaken is not open to dispute. Thus (1). If the information in a file is to be used in that sort of way, then the file may include a great deal of information in the file bearing on the global inferences that the thinker in possession of the file is in a position to make. This information will not be gained through any form of acquaintance with the individual the file is on, and perhaps not through acquaintance with any particular

individual at all. Thus (2). Given that things are not going wrong when this happens, there can be no acquaintance norm that gets violated. That implies that there is no particular way of getting information that the file is based on. Thus (3).

From this, it seems to follow that any expanded file is thereby detached. Expansion just is detachment. If that is right, then the move to deny the existence of encyclopedia entries, and instead claim that expanded files do all the work, is flawed. I stress that this reasoning is speculative, though it is reasonably compelling. In any case, the categorical point stands, and is enough to rebut the suggestion that Absence is not a problem. I conclude that there are good reasons to think that the indexical model explanation is insufficiently general, and so compares badly with any more general account.

The indexical model is only one way of providing an account of mental files that goes beyond the basic model, but no other sufficiently well-developed model has been provided in the literature, and it is not obvious how any alternative could avoid running into problems akin to Absence for the same reasons. Mental files are components of how a cognitive system processes information. What feature of files would provide information about the information regarding its subject matter that does not involve functional restrictions on that information? If there is no other feature, then even if it does not work quite as the indexical model does, no alternative robust model could avoid the same problem.

5.5.3 Summary

Indexical explanation seemingly works well in cases where it is plausible that the thoughts involve restricted ways of getting information. It does not work at all in cases where there is no restriction. Since these are the core cases of referential coordination, this means that the explanation is not sufficiently general. Much of the indexical model, as a general view concerning the nature of mental files, may be correct despite this problem. But it is a serious point against the existence of indexical files that it does not provide an explanation of referential coordination in core cases.

5.6 *Conclusion*

This chapter has examined the explanation of referential coordination by mental files. An explanation of this sort should be given in two stages: the introduction of mental files by

theoretical stipulation as what explains the paradigmatic inferential dispositions, and substantive claims about mental files that provides an indication of sameness of reference adequate for making those dispositions rational. This second stage can be carried out either with the basic model of mental files only or with some more robust model. The problem with basic model explanations is that they seem to demand cognitive capacities that make the explanation psychologically unrealistic. The problem with robust models is that they seem to demand functional restrictions on the information associated with files that limits the scope of the explanation and so makes it insufficiently general. The basic idea behind explaining referential coordination by mental file remains viable, however, despite these problems.

Chapter 6 - Coordination Functions Theory

6.1 Introduction

This chapter presents an explanation of referential coordination that satisfies the requirements and meets the explanatory desiderata.

The theory I propose is a reconfiguration of the sort of basic model explanation discussed briefly in §5.3, making use of the distinction between a mental file and its *predications*, by which I mean the associations between a mental file and bits of information, of which there can be several per mental file. Predications serve as the basis for the paradigmatic inferential dispositions that a thinker can make. Predications of the same file account for the thinker's disposition to put together the information by conjoining and generalising. I claim that predications of the same mental file will have normative functions that provide an indication of sameness of reference for the thoughts they subserve. These functions provide a version of the apparatus of defaults and defeaters familiar from entitlement epistemology: coordination functions provide a default indication of sameness of reference that is open to defeat by other information. Coordination functions can explain referential coordination in a way that is inferentially basic and adequately personal-level (§6.2).

Unlike other basic model explanations, no special cognitive capacities are required, other than the capacity to be sensitive to the coherence of one's inferential dispositions with the rest of one's stock of information, so the account is psychologically realistic. Unlike the indexical model explanation, it is not restricted to any particular sort of case, and so is fully general. Because no theory of mental content or vehicle is required, no danger of circularity or incompleteness in relating coordination functions to patterns of referential coordination will arise. These points make the explanation better than the competitors, and constitute the main argument in its favour (§6.3). The explanation makes use of psychological entities and properties, albeit of a minimal sort, and makes no use of any representational property more basic than reference. The normative functional properties that it does make use of are not properties of representations, either of their content or vehicle. In terms of the three questions, this explanation puts us here:

	Psychologism	Referentialism	Weight
Senses	No	No	Heavy
Typed representations	Yes	Yes	Heavy
Mental files	Yes	Yes/No	Heavy
Coordination functions	Yes	Yes	Light

A feature of the explanation is that it makes the rationality of the paradigmatic inferential dispositions out as a matter of coherence, rather than conclusive reasons. These latter points may be taken as a reason to reject the account, as it provides no support for more ambitious projects in epistemology or cognitive science. But I argue that we should accept this deflationary consequence, something I will expand upon further in the next chapter (§6.4).

6.2 Predications, Requirements, and Functions

This section presents the coordination functions explanation in detail. The general idea is to give a version of the two-part explanation by mental file that uses only basic mental files. It makes use of the distinction between a mental file and its predications, and the role of those predications in providing information that can be employed by a thinker. I will show how this provides an indication of sameness of reference in that it provides for an apparatus of defaults and defeaters familiar from entitlement epistemology. The explanation concludes by showing that this apparatus is something to which the thinker can be sensitive without extraordinary psychological capacities.

6.2.1 Outline

The two basic commitments of the explanation that I want to propose are as follows:

Coreference Requirements: Predications of a mental file (the associations between a mental file and bits of information) are subject to the requirement that they are about the same thing; these *coreference requirements* arises from the role of predications in the information-processing structure, the mechanism that serves processes of gaining,

retaining, and exploiting information, that enables the thinker to undertake her projects and achieve her goals.

Coordination Functions: A predication of a file (the *base* predication) can have the normative function of carrying information about the same thing as other predications of the same file (the *target* predications), i.e. have *coordination functions*.

The two basic commitments can be combined to give a third, which involves a specific kind of coordination function:

Requirement-Based Coordination Functions: If two predications are subject to a coreference requirement, then each predication is the base predication of a coordination function with the other as its target predication.

With these three elements in place, we can give an explanation of referential coordination as follows:

Coordination Functions Explanation: If (i) P_F and P_{F^*} are distinct predications of a mental file F , then (ii) P_F has the function of carrying information about the same thing as P_{F^*} , and P_{F^*} has the function of carrying information about the same thing as P_F ; (iii) these coordination functions are requirement-based; so if (iv) there is no information to indicate that P_F and P_{F^*} are about distinct things (no *defeating information*), then (v) the thinker has an indication that the information carried by P_F and P_{F^*} is about the same thing.

The key point in elaborating this will concern the rationale for the move from (i-iv) to (v).

If this explanation is adequate, then there is a straightforward extension to the case where there is an absence of referential coordination: thoughts are not referentially coordinated when there is an absence of a requirement-based coordination function. This points to a more nuanced understanding of Frege cases, since there are two possible scenarios for them to arise: the absence of referential coordination, and the presence of defeating

information. This points to a better understanding of the epistemology of identity, which will be covered when the explanation is extended to identity judgements (§7.3).

6.2.2 Predications and coreference requirements

The coordination functions explanation makes use of a distinction between mental files and their predications. Mental files, for the purposes of the explanation, are simply basic mental files. That is, they are mental particulars with file structure, which I earlier (§5.2) defined as what something has when it is associated with bits of information, organises that information, and has persistence conditions that do not depend on those associations (i.e. survives gain and loss of associations with bits of information). The notion of a predication is a simple extension of this idea: a predication of a mental file consists of a file and a bit of information, where the latter is associated with the former. It is what constitutes an entry of the file into the processes of gaining, storing, and exploiting information. It is important not to confuse a predication of a file with the bit of information; the use of the slightly awkward noun ‘predication’, as opposed to ‘predicate’, is meant to help keep this clear (compare Recanati 2012: 37-8).

Predications of a file can be thought of, and for the purpose of the explanation that appeals to them need only be thought of, as a compound of a mental file and another sort of mental particular that represents having a property. Since a mental file can have many predications, and since its various predications are distinct, making use of the notion of a predication puts the proposed explanation more in line with sense theory and typed representations theory; there are two (or more) entities involved, not just one. But where sense theory and typed representations theory had problems with saying how it is that these entities bear on the rationality of the paradigmatic inferences, things are a little more straightforward where file predications are concerned (§6.3.2 provides more by way of comparison).

The distinction between a mental file and its predications is an extension of the basic model, but only a minimal one, in the sense that it does not require files to have any properties that are not implicit in the basic model itself. Just by virtue of having file structure, a mental file is apt to have predications. Whether all mental files need to have information associated with them or not, and so will all have predications in this sense, I take no stand. We are only interested in cases where they do. The distinction is therefore only as controversial as the existence of mental files in the first place, and that is a matter

to be settled by explanatory considerations. Mental files are to be introduced as per the first stage of the two-stage explanation I suggested as the framework for mental file theories of referential coordination. The reconfiguration of mental file theory, making use of predications, takes place at the second stage. Rather than the mental files themselves being made the subject of substantive claims that bear on the rationality of the paradigmatic inferential dispositions, it is their predications.

With the distinction between files and their predications in hand, we can now state the substantive claims made by the coordination functions explanation. The first such claim is that predications are subject to coreference requirements.

Mental files are part of a mechanism employed in the service of some wider purpose or purposes. These purposes can be found in any aims that a thinker might have, the means of realising which are, in part, the topic of projects and intentions. The thought behind the idea that predications of a file are subject to coreference requirements is that some of these aims are, and all potentially will be, served by carrying out inferences that exploit the thinker's information. Information is gained, through various means, retained, in the various forms of memory retention, and exploited, in action and inference. The most basic way of exploiting information in inference is through the paradigmatic inference: one conjoins information one has about something together and existentially generalises about things in general. This makes available more information to be retained and exploited.

This sort of exploitation of the information through inference needs to be informative in a way that is manifestly reliable if it is to be exploited further in a rational way. This is the idea that I glossed in the first chapter as the inferential dispositions *making sense* to the thinker. This is meant to mark a distinct property from those dispositions involving inferences that are truth-preserving or valid. The idea of validity is obviously important for the purposes of spelling out coreference requirements, however. For our particular example, for a thinker to make use of the information that there is something with a certain syndrome of properties, that there is something that is a Roman orator, in going about her business in a way that makes sense to them, they need a sense of why they think that. More to the point, there must be something that points to it being *correct* information. An indication of the correctness of information is a minimal requirement for it to be apt to be relied upon in carrying out projects and plans.

Coreference requirements come from the fact that, when a thinker makes use of information gained through subsequent inferential exploitation of information she had previously, there is a requirement on the initial information so that the new information is gained in a way that makes it apt to be relied upon in carrying out her plans and projects. The inferences' being rational, in the relevant sense, means the thinker having a reason to treat the *new* information as coming from *old* information in a reliable way. When inferences provide new information that can further the projects and plans of the individual, the new information needs to be manifestly reliably correct in this sense.

This is hardly news, of course. The hope is that it is all fairly obvious and uncontroversial. It perhaps bears emphasising that these are meant to be claims about the rationality of information use, that is, about the reasons a thinker has for making the inferences she is disposed to make. The point of drawing attention to them is that it means something important for the explanation of referential coordination. Referential coordination concerns one, albeit central, means by which new information can be derived from old information in a reliable way.

For a thinker to make manifestly reliable use of the information that there is something with a certain collection of properties—that there is something that is both Roman and an orator, as in the example—in carrying out her projects and pursuing her aims, she needs to be sensitive to the basis on which that information was gained, and exercise that sensitivity. This is what an inference making sense to a thinker amounts to. The disposition to make such inferences is what matters here. The disposition to make use of the information that file predications contain in the paradigmatic form of inference is a disposition to get new information. Admittedly, the gap between the old and new information in this case is not very large. But it is important, because the disposition will be present whenever more interesting, more possibility-excluding and more action-guiding, inferential dispositions are present. And it is distinctive, because reliability in this instance requires the old information to be about one and the same thing.

To be reliable, gaining new information from old by conjoining and generalising requires sameness of reference. A file's having predications is what makes it enter into the mechanism that serves as the basis for this, and it is predications that are subject to coreference requirements. Because file predications and the disposition to exploit the information in the paradigmatic inferences go together, this means that all predications

of a mental file, when that file is part of a mechanism that serves the pursuit of projects and aims, are required to carry information about the same thing.

6.2.3 Requirement-based coordination functions

The second substantive claim is that coreference requirements impose coordination functions. To spell this out, it needs to be explained what a coordination function is, and then why coreference requirements can be said to put them in place.

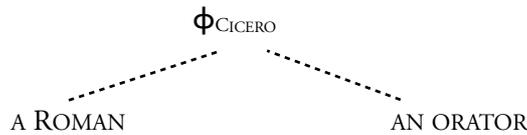
Coordination functions, in general, are properties of file predications that relate them to one another in terms of the identity of the thing about which they carry information. Coordination functions are *functions* in the sense that they concern what file predications are *supposed* to do. These are functions in the *normative* sense, as opposed to the descriptive sense that concerns only what something in fact does. Analogies that come easiest to hand concern biological entities, like bodily organs, and artefacts, like tools, or parts of machines. The heart has the normative function of pumping blood around the body; this is true even when a heart is in fact not pumping blood, in which case something has gone wrong and medical attention is required. Scissors have the normative function of cutting; this is true even when they are blunt, in which case the scissors are defective. A crank shaft has the function of converting the reciprocal movement of an engine's pistons to the rotational movement needed to move wheels forward; this is true even when the engine is idle.

Coordination functions, then, are like pumping functions, cutting functions, and movement conversion functions. They concern what things of a particular kind characteristically do as part of their being instances of that kind, and define a normative standard for them. File predications belonging to a mechanism that serves the pursuit of projects and aims characteristically carry information. There can be overlap in what that information carried is about. In some cases, there will be a standard of success or failure to the effect that things are going well when there is overlap, and badly otherwise. Saying what these cases are requires bringing in coreference requirements. To say that coordination functions can be *requirement-based* means that they obtain *because* there is a coreference requirement on two predications. The normative nature of these requirement-based coordination functions is a product of the contribution file predications make to cognitive processes that further a thinker's projects and plans. Things are going the way they should when the information they carry is about the same

thing because this enables reliability of new information gained through the conjoin-and-generalise inference.

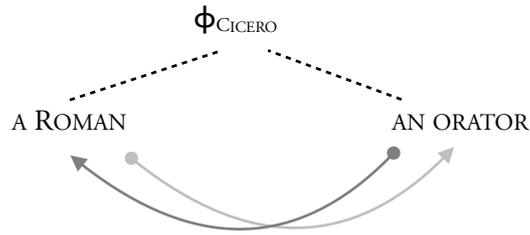
For a given coordination function, we can distinguish between its *base* and *target*: a predication that has a coordination function is the base predication, and any predication that the base predication has the function of coreferring with is the target predication. File predications can be both bases and targets. Where there is a requirement-based coordination function, all predications of the same file will be bases of coordination functions with all other predications as targets.

Let's apply this to our original example. Lucy thinks that Cicero is a Roman and that he is an orator, and these thoughts are referentially coordinated. This means that she is disposed to infer, from these thoughts, that there is something that has both of those properties. As per the two-stage explanation, we posit a mental file with two predications to account for this disposition. We can diagram this like so (compare the graph-theoretic representation of Frege cases given in Pryor forthcoming):



' ϕ_{CICERO} ' is a picture of the file; 'A ROMAN' and 'AN ORATOR' along with the lines connecting them to the file are pictures of the predications. A better representation would instead add a further predication, 'CALLED "CICERO"', rather than have it written into the file, but I omit it to simplify the diagram.

Because ϕ_{CICERO} is associated with the predications A ROMAN and AN ORATOR, Lucy is disposed to perform the paradigmatic inference, and because this inference provides information that may be relied upon in Lucy's pursuit of her plans and projects, this means that these predications are subject to a coreference requirement. Consequently, A ROMAN is the base for a coordination function that has AN ORATOR as its target; A ROMAN has the function of carrying information about whatever AN ORATOR does. AN ORATOR is the base for a coordination function that has A ROMAN as its target; AN ORATOR has the function of carrying information about whatever A ROMAN does. We could diagram that using directed lines, like so:



A more realistic representation of the example would involve a great many more predications and so more lines for all the coordination functions. Obviously that would be a very complex diagram. Hopefully the simplified picture gets the idea across.

File predications and their coordination functions can therefore be thought of as making up a network of normative functional relations based on coreference requirements. Thoughts that involve information belonging to the same network of such relations is, roughly, how we can think of what it is to exhibit sameness in ways of thinking. What puts this network in place is the contribution that inferential dispositions make to a rational life, and the grounding of those dispositions in the relevant file predications.

A worry that might be raised at this stage is that the explanation employs redundant elements. Coreference requirements and coordination functions may appear to be just the same thing given different labels. After all, a requirement on something, and a function of something, are both normative properties, and coreference requirements on file predications and coordination functions of file predications are both normative properties concerning the subject of the information that a file predication carries. Coordination functions will obtain whenever there is a coreference requirement on file predications. So it is true that they are very close. But they should be distinguished, for three reasons.

Firstly, the latter *explains* the former. The aetiology of the normative functions needs emphasising to make clear how they are based in the minimal information-processing architecture outlined above. Secondly, the idea of a coordination function articulates a form of rational explanation in a way that the idea of a coreference requirement does not. This form will be set out when we see how the presence of a requirement-based coordination function provides an indication of sameness of reference. Thirdly, there is also at least one other way of grounding coordination functions that supplies this form of

explanation, and this will be important for understanding the less basic way for identity to figure in thought, identity judgements, as I will argue in §7.3.

There are some related aspects of coordination functions that are worth being made explicit. One is that something can have a normative function to do something because it characteristically *does* do that thing. The point can carry over to explaining why that thing exists: that a sub-mechanism is the sort of thing that brings some state of affairs can explain its continuing presence in a mechanism. Simply being a predication of the same mental file is not sufficient, on this picture, for being the base of a requirement-based coordination function with other predications of that file as targets.

The other aspect is that the normativity of requirement-based coordination functions is *derivative*, in the sense that it is not basic from the standpoint of goal-oriented, or teleological, phenomena. There are various theories of such phenomena, including theories of normative functions, that deal with their metaphysics, particularly concerning the possibility of grounding them in non-teleological facts. For the purposes of the present discussion, it is not necessary to argue for or even assume such a theory. It serves to point out the following is an obvious possibility: something can have a normative function as a consequence of there being a requirement on it in virtue of its being a part of a wider mechanism that serves some wider purpose. The analogy with bodily organs illustrates this: because the heart pumps blood, and because it is part of a circulatory system that serves the purposes of the body, there is a biological requirement on it to pump blood.

Coreference requirements are something additional to the existence of mental files and their predications. One needs a whole thinker—a person, in a very full sense of that word—with a sensitivity to the reliability of her information and call for exercising that sensitivity. Mental files are only a part of a person. The idea that they have predications that are subject to coreference requirements brings in things that are presupposed by the puzzling phenomenon to be explained. In combination with the two aspects of requirement-based coordination functions just noted, this makes for a useful feature of any explanation that makes use of them: a commitment to requirement-based coordination functions is only committed to purposive cognitive processes that involve inferential processes that require sameness of reference. It is not committed to any particular view about what these cognitive processes must be like, nor to any view about the ultimate metaphysical nature or grounds of normative requirements.

This does not mean that nothing additional to basic mental files is brought in, of course. The two substantive commitments concern the existence of coreference requirements, and the imposition by these requirements of coordination functions. But the extent to which this involves moving beyond the basic model of mental files is minimal. The addition of requirement-based coordination functions is constrained by the particular explanatory use to which they are being put. This does not have anything to do with the functional properties mental files have as apart from or more basic than their playing that role. Nor does it impute to thinkers any special cognitive capacities with respect to their files that we have no reason to think they have. Granting for the sake of argument the existence of mental files, and the distinction between files and their predications, if one is interested in explaining referential coordination, then it should be relatively uncontroversial to hold that, if there are file predications, then they have requirement-based coordination functions.

6.2.4 Defaults, defeaters, and rational sensitivity

Let me briefly rehearse what is to be explained. Thinkers have thoughts that they treat as being about the same thing, and the mark of this is the presence of the paradigmatic inferential disposition to conjoin the properties ascribed to the things, and to quantify over things in general, to the effect that something has both of those properties together. This is not simply something that just happens to us, but rather something that makes sense to us. It is a *rational* disposition to have. As such, it stands as the basis of our possession of a coherent conception of things. This much is manifest. The difficulty is that it is not manifest what makes the disposition rational. The fact of the identity of the things thought about cannot suffice, since there are cases where that identity obtains but the rational disposition does not and could not.

In our example, Lucy thinks thoughts that we could report as her thinking that Cicero is a Roman, and that he is an orator. In one sort of case, her concluding that there is something that is a Roman orator need not be a rational move for her to make; she does not know that Cicero is the guy she is talking to, so it would only be rational if she had a belief in the identity of Cicero and the guy she is talking to. In another sort of case, she will be rational in making or being disposed to make the inference to the same conclusion without needing to have a belief in any identity. The difference must consist in the presence of an indication of sameness of reference that is absent in the other sort of

case. Sameness of sense, sameness of type of mental representation, the association of the information with the same mental file—explanations of these various sorts fail to provide such an indication in a satisfying way. Requirement-based coordination functions provide an explanation that does better. Here is how it works.

There are three components to the way in which requirement-based coordination functions connect with and so explain the rationality of the paradigmatic inferential dispositions. The first is the presence of a default presumption that something with a normative function is carrying it out. The second is the possibility and the absence of defeating information, where defeating information provides an indication that the normative function is not carrying it out. These components make up what I will call an *apparatus of defaults and defeaters*. This apparatus may be familiar from the epistemological study of entitlement. Entitlement is a species of warrant for, or a positive epistemic status of, belief. In Dretske's (2000) way of putting it, warrant is an epistemic *right* to a belief without a corresponding epistemic *duty*. The epistemic duty is to justify one's beliefs by providing evidence or reasons in support. As per the distinction that Dretske draws (Burge 1993b uses the same terminology in a parallel fashion; see Casullo 2007 for useful critical discussion), a belief to which one is entitled need not be justified. One may therefore have warrant for a belief without having evidence for it or reasons to take it as true. Nevertheless, the believer may still have warrant for that belief.

The evident advantage of an epistemology of entitlement is in the lack of inferential structure needed to supply warrant, thereby avoiding sceptical arguments. One example is the so-called Agrippan trilemma: the appeal to justifying reasons or evidence can always be questioned, and so justification must either go on indefinitely, stop arbitrarily, or go in a circle; in any case, the status of one's warrant for one's beliefs is in doubt. Entitlement, by contrast, is simply there in virtue of a belief being held by the kind of being that holds it, capable of forming beliefs in certain ways, in the kind of environment it is in. The apparatus of default and defeat works along the following general lines: by forming a belief in a certain way, there is a default presumption in favour of that belief being true. The warrant for that belief can lapse if there is information that indicates that that default presumption is false or unreliable, and warrant does lapse if the defeaters are themselves undefeated. Entitlement epistemology has been applied to most ways of coming to form beliefs, such as memory, testimony, and indeed inference (again, see Burge 1993b: 458 for a statement to this effect).

The idea that requirement-based coordination functions explain referential coordination is similar to this, though is not to be confused with this sort of appeal to the apparatus of defaults and defeaters. This is not least because the problem is not one of epistemic status, of rights and duties with respect to truth, but one of articulating the *reasons* that a thinker *has* (as opposed to, we might say, the reasons that there *are*) for a thinker to form beliefs, or extend their knowledge, or extrapolate from known information. The problem is therefore orthogonal to the one at which the idea of epistemic entitlement is aimed.

Nevertheless, it is interesting to note the similarity between the problem for warrant that gets posed by the Agrippan trilemma and the explanatory problem of referential coordination. In the former case, the problem is that any step in an inference one could make in support of a belief would itself be open to challenge; in the latter case, the problem is that any *additional* step in the inference one might bring in would itself be in need of referential coordination. This is what I earlier called the way in which referential coordination is *inferentially basic*. And there is nothing that makes articulating the structure of reasons had by a thinker using the apparatus of defaults and defeat impossible, despite it often being used to understand epistemic warrant in the absence of possession of reasons.

The first two components make up the indication of sameness of reference. The third component concerns how a thinker is sensitive to the presence of a requirement-based coordination functions. This is what makes the apparatus of defaults and defeaters work in supplying an indication of sameness of reference that is a reason that a thinker has. There must be some way in which the presence of a default and the absence of defeat is something to which the thinker can be sensitive. That this is indeed in place becomes clear when the way in which requirement-based coordination functions work is laid out.

The idea is that, when something has a normative function, this provides the default information that it is being fulfilled. It is subject to an expectation to the effect that it does something so long as things are going well, and so the default information can be relied upon in the absence of indication that things are not going well. In other words, something's operating under the expectation that it does something is itself an indication that things are going well and that it is doing it. So normative functions provide an indication that the function is being carried out. There can be a counter-indication to the effect what has the function is not carrying it out. In which case, the counter-indication overrides the indication provided by the normative function. This is the defeater.

Everything I have said so far about normative functions applies also to coordination functions. When two file predications are required to carry information about the same thing, there is an expectation that they do. Absent defeating information, there is an indication that they carry information about the same thing. Since referential coordination concerns the possession of information that one thing has two (or more) properties, defeating information will consist of information that two properties cannot be possessed together, or are sufficiently unlikely to be possessed together.

With this said about coordination functions in general, the same applies to requirement-based coordination functions in particular, which provide an indication of sameness of reference in exactly the same way. What is distinctive about them, or what makes them particularly suited to our purposes, is that the normative function imposed by a coreference requirement does not depend on the possession by the thinker of any additional information that bears on the identity of what the information is about. So this means that the indication of sameness of reference is inferentially basic. The default position, the expectation that file predications carry information about the same thing, is not something that needs to be explicitly formulated by the thinker.

How does defeating information work? The presence of a defeater will over-ride the default, and so render it irrational to perform the paradigmatic inference. But it does not remove the presence of the indication of sameness of reference provided by the default presumption, because it does not remove what put the requirement in place. It therefore exerts rational pressure to change *something*. One is doing well as a rational thinker if one is sensitive to the fact that a change needs to be made in such a case, and makes the right change. There are various forms that defeating information could take, and various ways that the result of any potential change would be a more coherent overall combination of information and inferential dispositions. We might distinguish between *putatively* defeating information, and *genuinely* defeating information. Putatively defeating information can take different forms. In the relatively uninteresting case, this will involve dropping the defeating information and maintaining the set of coordinated thoughts (*revoking*), or maintaining the defeating information and dropping either of the coordinated bits of information, so changing to different set of coordinated thoughts (*revising*). More interestingly, it may involve dropping the disposition and changing to a set of uncoordinated thoughts (*reconstructing*). This third kind is genuinely defeating information.

Cases where there is merely putatively defeating information, information that calls for revising and revoking, will not be hard to imagine (Lucy is told by an unimpeachable authority that oratory was not practiced by Romans; having met Cicero, she realises that Romans really did practice oratory). Cases of genuinely defeating information calling for reconstructing one's files, are more interesting and a bit trickier. The sort of informational states that would make dropping the inferential dispositions the rational thing to do would need to provide an indication that the identity required by the inference to be reliable is incorrect, even though the inference does not itself include a mistaken identity belief. What might this sort of information state be? In any realistic case, a specification of such an information state will be highly complex, as it would depend in various ways on more or less the whole of a thinker's information state. Potentially, any bit of information could (which is not to say, will) bear on this. But the following toy example shows the principle.

Take the following case:

Lucy sees a dog on each day $d_1 \dots d_n$, and calls it Fido. She believes various things about Fido (he is a dog, he is a labrador, he has fleas, she saw him on days $d_1 \dots d_n$), and as a result is rationally disposed to infer that there is something that is a dog, is a labrador, has fleas, that she has seen on days $d_1 \dots d_n$.

We can change the key detail to get a very similar case:

Lucy sees various dogs that look exactly alike, and is under the impression that there is one dog, whom she calls Fido, that she sees everyday; she comes to believe various things about him (he is a dog, he is a labrador, he has fleas, she saw him on days $d_1 \dots d_n$), and as a result is rationally disposed... etc.

These two cases are exactly alike in how we should think of them as per the coordination functions explanation, even though from the outside, as it were, quite different things are happening. The second case could be continued like so:

Lucy is told that one of the dogs she saw on day $d_1 \dots d_n$ was a different dog to all the others, but is not told which one.

Suppose that Lucy is told is true, and that there is no way she could rationally reject it, and suppose also that she saw no other dogs on those days (one gets a sense of how complex a real-life case will be). Now she has a defeating piece of information that means her inferential dispositions need to be changed. The default indication does not cohere with the rest of her information. As a rational thinker, she will be disposed to shift from having one FIDO-file to several files because of receiving that new information and being sensitive to how it fails to cohere with the identity required by the inferential disposition involving the FIDO-file. The question of how these files relate to each other, of whether she should form any identity judgements (about which more in §7.3 below), is then open.

It is important to stress that, just because reconstruction is possible in this way, the presence of the original mental file does not depend in any way on there being an identity judgement in place, or the thinker needing to *recall* any explicit acts of identification. The reconstruction involves creating new files, and dividing up the information in the way that makes the new information state coherent, and this need not be guided by any previous explicit beliefs in one thing being identical to another. Cases of genuinely defeating information will presumably be relatively rare (we had to stipulate that Lucy had no reason to rationally reject the new information). In fact, one would *expect* that to be the case, since otherwise the information processing system—files, predications, and the rest—would not be well designed. Presumably, if our environment was and had always been one where the potential for such mistakes was very common, then our cognitive lives and the mechanisms subserving them would need to be very different. The point is that there will always be the potential for putatively defeating information, and so the potential for genuinely defeating information too, and so the default is open to defeat and a matter of rational sensitivity.

How does the default presumption supported by a coordination function provide an indication of sameness of reference that is a reason the thinker *has*? The answer is that a requirement-based coordination function involves things to which a thinker capable of referential coordination must be sensitive, namely: the possession of bits of information, the disposition to put them together in inference grounded by the information being carried by predications of the same mental file, the fact that this disposition contributes

to the extension of her stock of information that may be used by her to further her aims and projects, and so the potential need for a change in her dispositions in response to other bits of information.

The indication of sameness of reference is adequately personal-level, since it involves a thinker's being sensitive to facts about her inferential dispositions. The explanation is also inferentially basic, because the operation of that sensitivity does not require bringing in any additional inferential steps concerning identity. Therefore, the requirements on an adequate explanation of referential coordination are met.

6.2.5 Files are inessential but indispensable

The objection might well be made that, if coordination functions provide an indication of coreference in this way, by means of a default presumption that information is about the same thing, then mental files and their predications are not really a necessary part of the explanation at all. The thought would be that, on the present showing, all one needs are the thoughts and the right kind of inferential dispositions, and everything else could be said file-free, without any commitment to sub-personal psychology.

It is true that mental files and their predications are certainly not *necessary* for the kind of explanation of referential coordination being proposed. One could, coherently, run a similar account file-free, as suggested. But there are reasons for insisting on including mental files as part of a *good*, and so potentially *best*, explanation.

One reason is that, since mental files can be introduced to do the work of explaining the relevant inferential dispositions, and since those dispositions put in place a coreference requirement that grounds the relevant properties, the coordination functions explanation is better off being formulated in terms of mental files and file predications. The file-free explanation would leave these dispositions dangling. Given that there may be multiple ways for the dispositions to come about, and given that these differences may make a difference to how exactly the explanation works, remaining silent on the basis of the dispositions would potentially hobble the explanation at the point where including files and their predications makes it complete.

Sceptics about this point are referred to §7.3 for a demonstration regarding identity judgements. But they may also be answered by noting the ubiquity of references to file-

like entities in the literature in philosophy, linguistics, and psychology on topics directly related to referential coordination. The obvious thing to draw from the ubiquity of the appeal to things that have file structure is that there is an explanatorily powerful idea in the offing. A clear and particularly germane instance of this is the appeal to file-like entities such as prototypes and exemplars in the psychology of categorisation, mentioned in §2.2.6, and the idea of object-files in theories of visual processing and perceptual attention. Also related is the appeal to files in dynamic approaches to formal semantics, and the use of mental files in the philosophical literature to explain facts about singular thought, empty and fictional reference, and so on, mentioned in §5.2. Though seeing it through is beyond the scope of this essay, I want to suggest that there is potential for a large degree of explanatory unification if we adopt the file-involving rather than file-free explanation.

Articulating what the presence of an undefeated coordination function involves can be done by bringing in a (very minimal) theory about what goes on at the sub-personal level. This is what mental file theory supplies, and it does so in a way that has considerable benefits. One can only say so much at this stage, but it is not nothing. The file-free account is not obviously better, so this objection has little force. A reasonable conjecture is that, while additions or modifications to the model of mental files may be needed once the account interacts with explanatory problems elsewhere, these will likely preserve the basic idea that mental files and their predications sustain coordination functions sufficient for explaining referential coordination.

A further point against the file-free version is that it is not clear how to make thoughts themselves bases and targets of coordination functions without some account of what thoughts are. So although one could put the explanation in terms of the thoughts in question, rather than the file predications (and so dispense with mental files as well), this would bring with it a need for a theory of mental content or the vehicles of mental content, and this would be to presuppose something that should be answerable to a theory of referential coordination. Moreover, we have seen problems with theories of both kinds. One can account for the rationality of thought while remaining silent on the question of what thoughts are (to reiterate: mental files are not, *pace* Fine, bundles of thoughts; see §5.2). What matters is that the thoughts in question involve information we have about the world. A theory of how that information is stored and processed makes available an explanation of why it is rational to treat one's thoughts as being about

the same thing. So though not necessary, making use of mental files (more exactly, of *basic* mental files) presents a path of least explanatory resistance.

It should also be noted, however, that coordination functions theory is not inimical to the idea that thought involves mental representations. A (basic) mental file is a mental entity that is associated with and serves to organise information. It is a functional entity, in the sense that its nature is exhausted by this functional characterisation. There might be several ways to implement file structure. Mental representations can be considered one sort of implementation. Suppose a thinker has these Mentalese sentences in her ‘belief box’:

- | |
|---|
| <ol style="list-style-type: none"> 1. CICERO IS A ROMAN 2. CICERO IS AN ORATOR 3. THERE IS SOMETHING THAT IS
BOTH A ROMAN AND AN ORATOR 4. TULLY IS NOT A ROMAN |
|---|

The subject-place concepts in beliefs 1 and 2 are type-identical. All this needs to mean is that they are treated as being about the same, so that the predicates with which they are concatenated (and so are equivalent to file predications) brings about belief 3 being in the belief box. The subject place concept in belief 4 is type-distinct. This means that it is not treated as being about the same. This is a way of implementing a system of mental files. The file is implemented at a more fine-grained level by distinct though type-identical concepts (this is in line with Millikan’s (2000) thought about the functional equivalence of duplicates markers and dot markers discussed in §4.7 above).

Friends of mental representations would be advised to take up this option, as it removes the need to provide any more substantive typology to explain referential coordination. Type-identities relevant to referential coordination are just those that underwrite the paradigmatic inferential disposition. The work of providing an indication of sameness of reference is done at the functional level of the mental file, and the coordination functions explanation can then be run as above (see §7.2 for further discussion).

The coordination functions explanation therefore makes indispensable use of both the personal and sub-personal levels of psychological description. The point is that this does not mean abandoning the personal-level and ‘changing the subject’ in any problematic

fashion. In this way, the explanation *bridges* the two levels. Rationality at the personal level has a basis in a mechanism at the sub-personal level. I will expand upon this in the next section.

6.2.6 Summary

Referential coordination can be explained by amending the basic model to include the distinction between mental files and their predications. Predications of the same file are required to carry information on the same thing because of the role they have in grounding the paradigm inferential dispositions that makes them bear on the fulfilment of the thinkers projects and goals. This requirement makes them the bases for coordination functions with all the other predications as targets. The coordination function linking two predications indicates that they are about the same thing. This indication is open to defeat by information that indicates the opposite. Coordination functions therefore exert rational pressure on a thinker's inferential dispositions that potentially requires a change to be made. In this way, there is an indication of sameness of reference that supplies a reason for the thinker that is inferentially basic and adequately personal-level.

6.3 Evaluation

In §1.6, I set out five desiderata for an explanation of referential coordination. The explanation to be preferred is:

- (1) sufficiently general,
- (2) non-circular, and
- (3) psychologically realistic, in a way that
- (4) can be extended to providing a plausible account of identity judgements, and
- (5) is consistent with general facts concerning special ways of thinking.

The latter two desiderata have to do with the way in which an explanation of referential coordination interacts with other problems of ways of thinking; I will deal with them in the next chapter. In this section, I will argue that the coordination functions meets the first three without difficulty, and provide comparisons with the competing explanations to show that it should be taken as the best explanation.

6.3.1 Meeting the desiderata

Taking each desideratum in turn:

Generality: The coordination functions explanation has no problem with generality, since the explanatory apparatus applies without reference to any particular features that might differ across instances of the paradigmatic inferential dispositions. They are not limited to cases where there is a particular way of getting information from something, since the means by which the information associated with the file was gathered does not appear in the explanation of referential coordination. In particular instances, there may be defeating information that makes reference to some point about where the bits of information came from. But this presents no problem of generality for the explanation.

Circularity: Similarly, the coordination functions explanation does not have problems of the sort that undermined sense theory, since none of the entities or properties that it relies on are themselves to be cashed out in terms of particular patterns of inference. No *a priori* reflection directed at formulating reference rules or defining epistemic intensions is needed to determine the way in which the thought represents its object. Reference is treated as a primitive representational property, so no explanatory relationship between rational and representational facts enters into the account.

An objection might be made at this point about the role that the notion of a cognitive life, made up in part of a thinker's rational conception of things and her pursuit of projects and aims, is playing in the coordination functions explanation. It must be admitted that it does play a role, and so is part of the explanation: referential coordination is explained by coordination functions providing a default expectation that reference of thoughts is available for defeat, so making the coherence of one's inferential dispositions with one's broader stock of information available to the thinker's rational sensitivity. That is, it is a matter for her rational sensitivity that her inferential dispositions hang together with what other information she has, and if they do not, then to the extent that she is rational, she will be disposed to modify something about her informational set-up. This presupposes that a thinker has such a rational sensitivity, and so is capable of having a cognitive life characterised by the possession (and so the maintenance) of a rational conception of things. So much can be taken as given for the sake of argument. But, the objection goes, referential coordination is an interesting topic

for explanatory theorising *because* of the role it plays in structuring our cognitive lives by making the possession of a rational conception of things possible. So the explanation is viciously circular. It takes for granted what needs to be explained.

In response, I simply deny that this is the case. It is true that referential coordination makes having a cognitive life characterised by the possession of a rational conception of things possible. And it is true that having a rational sensitivity, being sensitive to the coherence of one's stock information, is presupposed in the coordination functions explanation. These points do not mean that the coordination functions explanation is in any way viciously circular. Referential coordination is a fundamental feature of thought for having a rational conception of things, but it is only a part of what it takes to have one. The objection, that the coordination functions explanation is circular, effectively assumes that any non-circular explanation must be such as to allow us to comprehend a whole cognitive life by explaining its parts without reference to that whole, and this assumption is unwarranted. The explanation of how a part of something contributes to the workings of the whole can, and in some cases perhaps must, make use of its place in the whole, and so help itself to the idea that there is an adequate enough prior understanding of that whole. There is no question of an explanation by *reduction* here (those sympathetic to anti-explanationism may perhaps find this congenial, though it lends no support to that position).

The kind of rational sensitivity that the explanation uses is presupposed in the way that the explanatory problem was set up. So it is legitimate to appeal to that rational sensitivity as part of the explanation, and legitimate to state what the rational sensitivity involves by reference to what was presupposed in the explanatory problem. To put things somewhat picturesquely, the problem can be restated as the need to explain how sameness of reference can be a topic for our rational sensitivity, given that it cannot always consist in an explicit judgement of identity. The coordination functions explanation is not circular in the way it does this. So the objection can be resisted, though it does raise an interesting facet of how the explanation works. And it remains the case that rational sensitivity, in the sense I have been using that term, is itself something in need of explanation.

Psychological realism: The coordination functions explanation is committed to properties of mental file predications that are based on coreference requirements. Though this may seem like a relatively extensive set of commitments, the explanation is really fairly

conservative. Nothing is introduced that we have no reason to think is there anyway just in setting up the explanatory problem, other than the theoretically stipulated basic mental files. In addition, no demands are made on thinkers other than to have the sorts rational dispositions that we are interested in in the first place. The coordination functions explanation simply brings out what was, in a sense, already there in the data that prompted the problem of referential coordination, with a minimum of theoretical novelty.

In this connection, commenting on Recanati's comparatively robust indexical model of mental files, Hall writes:

According to an opposing paradigm, the primary function of files is simply to group information together in cognition. A special case of this occurs when that information is taken to concern the same external object, but in general the information stored in files needn't even be grouped together by subject matter. If Recanati is to maintain that the acquaintance-based function of his mental files is somehow primary, he needs to offer more by way of argument for that claim, especially if he thinks it has certain empirical consequences for the evolutionary relationship between acquaintance-based files and files with derived functions. (2013: 19)

Hall's 'opposing paradigm' might as well be the basic model; taken that way, the point seems to me correct, and coordination functions theory is perfectly in line with it. We have as good a reason to think that mental files have predications with requirement-based coordination functions as we do for thinking that there are mental files in the first place. And we can go further. The introduction of mental files was supposed to be a sort of temporary hypothesis, to be fully justified by the explanatory work that it can do. The fact that the basic model of mental files is retained while explaining both the dispositional and rational parts of the paradigmatic inferential behaviour means that the explanatory advantages of the explanation carry over to mental file theory more generally. A virtuous circle, if ever there was one.

6.3.2 Comparisons

Comparison with the other explanations discussed so far shows that and why the problems they faced do not arise for the coordination functions explanation.

The coordination functions explanation contrasts most sharply with sense-theoretic explanation. There is no additional notion of content needed, and certainly no range of entities individuated by their role in rational inference. Unlike sense theory, therefore, it is no part of the coordination functions explanation to provide a way of relating a thinker's cognitive capacities to these special entities. The problems of explanatory circularity and incompleteness that that presented do not arise. In place of a theory of grasp, as I called such a theory, there is the theoretical introduction of mental files and their predications. The relationship between this and a thinker's thoughts and cognitive capacities is a matter of cause, interaction, and consequent normative function; there is no analogue of the handles challenge.

It might be thought possible that coordination functions provide a kind of back-door for the introduction of sense theory, perhaps serving as a way of overcoming the handles challenge. This would be a mistake, however. The sense-theoretic explanation works by attributing sameness of sense to referentially coordinated thoughts, which constitutes an equivalence between thoughts. But there is no reason for thinking that referential coordination involves an equivalence between thoughts that would cause problems for the coordination functions explanation. Insofar as coordination functions theory does introduce a kind of equivalence, it is one that is dependent on a thinker's inferential behaviour, or disposition to engage in inferential behaviour, with respect to her thoughts, and does not demand explanation by the presence of an equivalence relation at some deeper explanatory level. To be treated *as* the same does not presuppose that something *is* the same in some further way, and nor does it require explanation by positing some deeper equivalence.

Relations of sameness and difference in ways of thinking are primarily to be found in the patterns of inferences that the thinker is rationally disposed to make. The sense theorist deals with the problem of explaining these patterns by *reifying* ways of thinking, in the form of senses. She then supposes that there is an interesting explanatory relationship between these reified ways of thinking and the inferences involving the thoughts in question. Quite apart from the fact that this explanatory move faces problems with circularity, ways of thinking need not, and should not, be reified in this way, and the coordination functions explanation does not do so.

It perhaps bears repeating that sameness of mental file does *not* suffice for sameness in way of thinking (that is, for referentially coordinated thoughts). A lot more is required, including coherence with the thinker's broader stock of information. This reflects the fact that the rationality of the paradigmatic inferential dispositions need only amount to it being coherent for the thinker to treat her thoughts as being about the same thing, rather than there being a once-and-for-all demonstration that they are about the same thing. The sense-theoretic explanation, by contrast, imposes this stronger condition. This marks a difference in approach to the problem that is really fundamental where the comparison with sense theory is concerned. I will go into this in more detail below in §6.4.2.

Moving on from sense theory, the contrast between coordination functions theory and typed representations theories is, perhaps, less pronounced because of the shared referentialist and psychologicist character. But the fact that they necessarily make use of a theory of mental representation, and are obliged to say why their favoured typology should match patterns of referential coordination between those thoughts in an explanatorily satisfying way, makes it no less acute. Coordination functions theory faces no such obligation. Again, there should be no question of coordination functions providing a back-door to a typed representations-theoretic explanation. Though sets of target-base pairs can be seen to constitute a kind of typology, it is not itself a typology of mental representations, at least not in the relevant sense. The way in which typed representations can implement the mental file part of the coordination functions explanation shows that typed representations might play some role in the best overall theory, but the typology will not be what explains referential coordination.

The fact that the coordination functions explanation bridges the personal and sub-personal levels of explanation provides a further explanatory advantage over typed representations theory. Problems of psychological realism that we saw in the case of the originalist and neo-empiricist typologies cannot arise. There is no need for extraordinary capacities of introspection or extraordinary sensitivity to the mechanisms subserving our information gathering and processing capacities. As I have already mentioned, the theory remains neutral on questions as to the nature of thought, especially as to whether it involves mental representations or not, and if it does, whether they are amodal, perceptual, sentential or non-sentential in format. This is all consistent with there being mental representations, and with their being typed in a way that maps on to patterns of referential coordination, but the explanation makes no commitment on either point.

The contrast between the coordination functions explanation and the sorts of explanations of referential coordination to be gleaned from extant treatments of mental file theory will be less pronounced still. While the contrast between the coordination functions explanation and the indexical model explanation is pronounced, there are some clear similarities with the two basic model explanations that I briefly examined in §5.3. This is particularly true of the account drawn from Lawlor's (2001) proposal concerning file maintenance dispositions. This proposal runs very close to what I said earlier about a thinker being sensitive to the coherence of her paradigmatic inferential dispositions. In some ways, Lawlor's treatment and the coordination functions explanation share several elements. They differ in how these elements are put together, and these differences are important.

One difference between Lawlor's account and the one defended here is that she takes files to be constituted by the presence of file maintenance dispositions (op. cit.: 79-91), which dispositions are exercised in the course of carrying out 'reidentificatory aims' (op. cit.: 75-79), whereas on the present account, files are mental entities that bring about the paradigmatic inferential disposition, and they have predications that are subject to coreference requirements in virtue of properties of the thinker that are independent of those dispositions. The dispositions to revise, revoke, and reconstruct are contingent on mental files subserving the dispositions of a thinker who is rational in a broader sense, rather than essentially tied to the existence of the file as such. Though important for the character of the two explanations, this difference can be set aside.

The more crucial difference between the two accounts can be summarised by saying that, for Lawlor, they constitute an *upstream* attempt to maintain an intentional relation (op. cit.: 72-5), whereas on the coordination functions account, they constitute a *downstream* attempt to maintain a coherent state of information. Lawlor's file maintenance dispositions include recognition and reidentification, which are upstream in the sense that they concern how information gets associated with a mental file. These bear on referential coordination because they make that file intentionally related to something, so the information is information about the same thing. On the coordination functions explanation, the parallel dispositions (to revise, revoke, or reconstruct) are downstream in the sense that they deal in information associated with the file with no reference to where it comes from, and involve other bits of information not associated with the file. Coherence kicks in because of the presence of the paradigmatic inferential disposition

and the need to make changes as a result of it. Both accounts therefore tie referential coordination to dispositions to make changes to one's files, but in quite different ways.

The problem with Lawlor's account was that appealing to file maintenance dispositions involves the explanation in some loss of psychological realism (we do not in general have or need to have extensive awareness of the epistemic status of our information in order to have referentially coordinated thoughts). This is not the case for the dispositions involved in the coordination functions explanation. These demand much less than being sensitive to what bears on the epistemic properties of one's information, including things perhaps in the irretrievable past. Referential coordination can be a feature of a thinker's thoughts even if she is lax in her epistemic self-monitoring. Coordination functions involve only what a thinker *must* have if she is to be capable of referentially coordinated thought at all.

6.3.3 Summary

The coordination functions explanation of referential coordination meets the first three desiderata where the competing explanations fail in various ways. In particular, the explanation avoids the pitfalls of alternative treatments of mental files, and so can preserve the comparative benefits of that approach in a satisfying way. This provides a strong argument to adopt it as the best explanation going.

6.4 *The Character of the Coordination Functions Explanation*

If we do adopt the coordination functions explanation, then what view of thought does it supply? I will begin by showing how it answers the three questions set out in §1.4. It will be seen that it supplies a lightweight, psychologistic, and referentialist view, and that as a consequence it provides a deflationary understanding of referential coordination that speaks against more ambitious philosophical claims that might be thought to be supported by reflection on rational inference.

6.4.1 Lightweight psychologistic referentialism

The coordination functions theory makes no use of a representational property more basic than reference (referentialist); it does make use of psychological properties and entities (psychologistic); and it does not require there to be any substantial relationship between what explains referential coordination of thoughts and what explains the

referential content of thoughts, or in other words, it makes no use of either side of the content/vehicle distinction as applied to thought (lightweight). As for the first two points, I take them to be obvious: coordination functions are properties of file predications, which are psychological entities, and their role in the theory does not require them to determine reference in any sense. I want to concentrate on the last point, though it may be obvious how the first two work towards supporting it.

The coordination functions explanation is lightweight because no part of the explanation bears on the nature of mental content, or on whether there are mental representations and if so, what they are like. That is why it makes no use of the content/vehicle distinction. It could be objected that, while one should grant as obvious that no use is made of the *content* side of the content/vehicle distinction, nevertheless use *is* made of the vehicle side, and so the view is not really lightweight at all. Why think this? On the one hand, it is common to see mental files spoken of as being files 'on' something, as having *topics*, as we might call it. On the other hand, there is the fact that file predications are said to be subject to normative requirements concerning the information that they carry. Having a topic and carrying information seem to be intentional properties. So there might appear to be a problem with seeing the view as lightweight.

That there is no problem can be shown by first dispensing with the idea that files have a topic, or rather, with showing how it can be accommodated. There seems to be something highly plausible about this idea. It is natural to suppose that the idea behind mental file theory must be that there is a mental file that is on the reference of the thoughts, and that their having that reference in common goes hand in hand with having the file in common, and the file being on that referent. It is natural to suppose that mental files are intentional vehicles of a sort. If it is granted that mental files have a topic, then their having a topic is a sort of intentional property. It answers to the way that intentionality is normally introduced: something has an intentional property if it is *about* something else, and having a topic seems to be a way of being about something else.

Mental files have intentional properties, if we grant the natural way of talking to have this consequence. But if they do, they do not do so in just the same way that thoughts are intentional. There is only a problem if we ignore the following point: the content/vehicle distinction that heavyweight explanations make use of is a distinction applied to *thoughts*. Mental files are of course not contents of thoughts, but nor are they vehicles of thought, at least so far as the basic model that is employed by the coordination functions

explanation goes. The intentionality of files, whatever exactly that comes to, does not mean that the coordination functions explanation is heavyweight.

The difference can be shown by noting a difference between the roles played by vehicles of thought and mental files. To be a vehicle of thought, at a very general approximation, is to function as a representation in whatever way is necessary for the content of that representation to have the distinctive role in a thinker's life that thought contents have—to be the content of beliefs, to have logical implications that are brought out in inference, to be expressed in attempts at communication, to be preserved in memory, and so on. That formulation is banal enough to be acceptable to all, or is meant to be. The tricky part is to say exactly what it might mean, in particular to say what those functions are exactly. But even so, it is clear that file predications do not fit into the representation slot of that formulation. Mental files do not function in that way. They are mental entities that *enable* some of those functions only. This goes for file predications in just the same way. Although they carry information, by being associated with a mental file, it is no part of the theory that makes use of them that they have the psychological role of thoughts. This remains the case even if, as was discussed at §6.2.5, mental files are implemented at a different functional level by things that *do* function as representations.

Put simply, the point is that no theory of mental representation is or needs to be upheld in the theory of referential coordination that I am recommending. This is a strength of the explanation, insofar as it represents the path of least explanatory resistance.

A question may remain about how the intentionality of mental files and their predications should be understood, if not as vehicles of thoughts. One way, of course, is to take them to be robustly intentional in the way that thoughts and sentences are normally taken to be. This attitude can be seen in Prinz's claims concerning the intentionality of long-term memory networks (see his 2002: 249-260). But given the difference I have been insisting upon, this option may seem unavailable, or at least unattractive. We might instead want to say that mental files are not used as representations in a way sufficient to make them vehicles of thoughts, but nevertheless have their uses, and have topics because of these uses.

There are a couple of ways of elaborating on this idea. One is to say that files having a topic is a way of keeping track of facts about how the cognitive system interacts with things in the wider world; being 'on' something is a handle for the mental file theorist, a

sort of courtesy term offered in recognition of the important work that mental files do that adds no more robustly intentional facts into the world. A different elaboration is to view files as components of a sort of informational model, a model that mirrors the structure of information (or misinformation) that a thinker has about the world, the individuals that inhabit it and the properties that they have; being 'on' something is a product of interpretation. File topics are something that we, as theorists of the mind, can interpret according to how it maps on to the world that it models. This makes the intentional properties of files more robust than being merely by courtesy, at any rate as robust as the interpretation that we can place on them.

The idea that mental files have their uses in this way should not be confused with how we interact with *non*-mental files. These are objects of purposive actions (searching for and through them, organising them, placing documents inside, and so on), and serve as tools for keeping track of information about things. Mental files are not tools used by the thinker in this way. As I emphasised above, they are elements in the cognitive mechanism at the sub-personal level. They thus clearly have an information gathering function, but it is not a function that consists in their being used, as tools in purposive action, as non-mental files are. But they do have such a use for the theorist, and this theoretical use can, perhaps, sustain intentional properties despite their not being vehicles of thought.

All three views on file topics are consistent with maintaining that an explanation of referential coordination that makes use of mental files can be lightweight. So the coordination functions explanation, as a lightweight explanation, can happily allow that files have topics. It also, perhaps, represents the beginnings of a novel path to conciliation between the different positions on the problem of intentionality taken by interpretationists and more straightforward realists. One could accept an interpretative view of intentionality at the level of mental files while also holding to a strongly realistic view of intentionality at the level of a person's thoughts: intentional facts at the sub-personal level are interpretational, and pertain to the grounds of non-interpretative intentional facts at the personal level. For the sake of completeness, I note that it is also consistent with eliminativism about intentionality, of the sort that is meant to be found at the level of thought, though this would remove the motivation for the theory, which was precisely to explain a feature of thoughts. Going into these points any further would take us well beyond the remit of the present essay, so I leave the topic there.

6.4.2 Coherence and connectedness

In the comparison with sense theory in §6.3.2, I said that the rationality of the paradigmatic inferential dispositions need only amount to it being coherent for the thinker to treat her thoughts as being about the same, rather than a once-and-for-all demonstration that they are about the same, and commented that this marks a fundamental difference between the two theories. I want to expand on that idea by way of elaborating on the character of coordination functions theory.

Any explanation of the rationality of the paradigmatic inferential disposition to conjoin and generalise that marks referentially coordinated thoughts will come with a view of what we might call the *mode* of rationality that they exhibit. By modes of rationality, I mean to pick out the different ways for the reasons that bear on inferential dispositions to be structured. Rationality in general can consist of different forms in which reasons make something rational. Reasons can be such as to decisively *exclude* certain possibilities. They can also be such as to *allow* some possibility or possibilities to be entertained. Accordingly, we get at least two different ways for the paradigmatic inferential dispositions to be rational. They may be so because there is something that conclusively demonstrates, or anyway appears to, that reference is the same, so as to exclude the possibility it is not; call this the *exclusionary* mode of rationality. On the whole, this is the mode of rationality that the explanations criticised in previous chapters go in for. Heavyweight explanations have the result that whatever indicates sameness of reference does so in virtue of being linked to what makes it the case that reference is in fact the same. So the indication of sameness of reference they provide is a reason in the exclusionary mode.

By contrast, the coordination functions explanation makes the rationality of the inferential dispositions belong to the *coherence* mode of rationality. Coordination functions, as properties of file predications, indicate sameness of reference by making the coherence of one's inferential dispositions with one's broader stock of information a matter of rational concern, something to which a thinker with the right capacities can be sensitive. Being sensitive will manifest in a disposition to make changes in one's inferential dispositions in response to putatively defeating information in the three ways mentioned above (§6.2.4). To reiterate a point made above, this explanation bridges what is going on at the personal-level and at the sub-personal level: what is coherent or incoherent is something that lies at the personal-level; what accounts for that coherence

being at issue at the personal-level lies at the sub-personal level. This bridging of the two levels makes available the lightweight nature of the explanation, and, consequently, the fact it belongs to the coherence rather than exclusionary mode of rationality. In light of the evident advantages that this sort of explanation has over the competitor explanations, the bridging of the two levels is a crucial feature of the account.

It is important, in saying all of this by way of characterising the explanation, to keep in view exactly what it is that was to be explained. Insofar as referential coordination, as a sort of inverse of what goes on in a Frege case, has received philosophical attention, it is largely under the guise of coreference *de jure*. I want to reemphasise the point that referential coordination and coreference *de jure* are distinct, because the characteristics of the coordination functions explanation of referential coordination will, at least at first appearance, be simply ruled out if what is being sought is an explanation of the stronger condition.

The coherence mode of rationality exhibited by the indication of sameness of reference provided by coordination functions does not look adequate as an account of the epistemic side of coreference *de jure*. The presence of a coordination function does nothing to secure anything along the lines of sameness of reference as a representational requirement. Whether or not something like the conjunction of representational and epistemic properties discussed under the rubric of coreference *de jure* really is exhibited in thought (or, for that matter, the conjunction of semantic and epistemic properties in language) must be a moot point. As far as understanding how identity can figure in an inferentially basic way, we can at least say that this notion does no useful work for us, regardless of the fact that some theorists have taken it as a starting point in their proposals about the nature of thought.

There is a revealing point of contact between this and how the basic model treatments of mental files owing to Lawlor and Schroeter are framed. In the course of their discussions of mental files, both direct their sights initially towards accounts of the stronger condition, combining epistemic and semantic/representational properties. Lawlor, following a discussion of Campbell's (1987/88) treatment of trading on sameness of reference in terms of sameness of sense, proposes to account for apparent sameness of sense. Since sameness of sense requires sameness of reference, this amounts to the same as apparent coreference *de jure*. Schroeter frames the inquiry into "the first person perspective of the individual on what's picked out by her own words and

thoughts” (2012: 179) as an inquiry into “apparent *de jure* sameness: the subjective appearance of obvious, incontrovertible and epistemically basic sameness of subject matter.” (op. cit.: 182)

Both Lawlor and Schroeter therefore frame their discussion in terms of the stronger condition plus a qualification so as to make it weaker: *apparent* coreference *de jure*, *apparent* sameness of sense. This qualification (which is not really commented upon by Schroeter, but plays a more important role in Lawlor’s discussion of warranted inference) puts the topic of investigation much more closely in line with referential coordination, in the sense that what is to be explained is a condition weaker than coreference *de jure*. Putting matters in terms of the stronger condition plus qualification does very little except express some negotiable assumptions about the nature of thought, and combine these with a more realistic view of what is actually required. Instead of approaching the phenomenon as we find it in our thinking in terms of a qualification on the stronger condition, coreference *de jure* or sameness of sense, my suggestion is that the stronger condition be simply swapped out for the weaker condition of referential coordination, so dropping the qualification with no loss of descriptive adequacy and no shortage of gain in clarity.

The point can be reinforced by noting a further overlap between the coherence mode of rationality adopted by the coordination functions explanation and what Schroeter calls ‘connectedness’ models of apparent coreference *de jure*. A connectedness model “appeals directly to the relationship between different token representational states established by the subject’s psychological dispositions.” (op. cit.: 191) The contrast is with ‘resemblance’ models, which Schroeter gives the general gloss of the idea that “[t]wo token representations are held to express the same meaning only if they resemble each other in a specific respect: each token must be associated with the very same pattern of understanding.” (op. cit.: 187; this distinction also finds expression in Fine 2007: 2-3) An adequate model, on Schroeter’s view, needs to account for accessibility (“the appearance of *de jure* sameness should afford reliable access to sameness of subject matter,” op. cit.: 184) and should allow for flexibility (“must be flexible enough to allow for the kind of variation in understanding we tolerate in rational inquiry and debate,” op. cit.: 186). Schroeter’s argument is that resemblance models cannot meet both demands at the same time, whereas a connectedness model can (op. cit. 189).

The lightweight explanation of referential coordination provided by coordination functions can be seen as a version of or variation on the ‘connectedness’ model that Schroeter recommends. It depends on the dispositions a thinker has with respect to her mental files and the coherence of those dispositions with her broader store of information, rather than resemblance in content or vehicle. The argument for this does not depend on bringing in considerations of flexibility, but rather in showing that there are explanatory deficits that come with cashing out accessibility in resemblance terms. The advantages of views of thought that support a connectedness model over a resemblance model are much clearer when put forward as a way of explaining referential coordination, rather than apparent *de jure* coreference.

Coreference *de jure* without qualification is naturally seen as a component of a rationalist conception of inference, and epistemology more generally. If rational thought requires *de jure* coreference, and if coreference *de jure* involves immediate and incontrovertible access to the fact that two thoughts refer to the same thing, then there is a sort of epistemic foundation, or at least epistemic ingredient, to our conception of things that is not up for confirmation or disconfirmation by empirical evidence. Not surprisingly, the coherence mode of rationality adopted by the coordination functions explanation of referential coordination does nothing to support this kind of epistemology, and it provides reasons for rejecting this way into it. This presents one strand of the deflationary character of the explanation, the consequences of which I will go into further in the next chapter.

There is no good reason, once referential coordination is made explicit as a distinct condition, to think that coreference *de jure* is a feature of thought. One would perhaps need to argue that, as it is a feature of language, it must be a feature of thought also. There can only be any philosophical pressure to treat thought and language as being on all fours in this way on a view of mental content that sees an essential role for something akin to the Fregean notion of sense. The explicitly articulated rationalist ambitions of sense theorists (see Chalmers 2012: xvii and Peacocke 2008: 150-159; also Peacocke 2003; for a view of Frege himself as a rationalist, see Burge 2013) may be laudable, but they find no support in considerations concerning the ways in which identity figures in inference. We may ideally want a kind of certainty, a ‘cognitive home’ in Williamson’s (1996) phrase, that can be had based on our rational capacities alone. But we may have to make do with coherence based on our fallible information-gathering capacities, with all the uncertainty and complexity that that entails.

6.4.3 Summary

As a lightweight explanation, the coordination functions explanation views referential coordination on the lines of the coherence rather than demonstrative mode of rationality, and does not suffice to support the stronger epistemic-cum-semantic notion of coreference *de jure*, or rationalistic ambitions in epistemology.

Combined with the earlier point about the lack of a theory of mental presentation or content, these points give content to what I meant, when introducing the coordination functions explanation, in saying that it has a *deflationary* character. On this view, referential coordination requires very little by way of theoretical novelty or location in broader philosophical ambitions. When one is disposed to treat one's thoughts as referring to the same thing, one is, with the addition of some fairly straightforward and theoretically minimal conditions, rational in doing so. The apparatus of defaults and defeaters provided by file predications, coreference requirements on file predications, and requirement-based coordination functions based on those requirements, is simply a way of saying what these minimal conditions are. Far from giving rise to a need for a theory of mental content, or a theory of mental representation, with the epistemological or psychological commitments these inevitably bring in, referential coordination requires only just enough theory to give an explanation, no less and certainly no more.

6.5 Conclusion

Coordination functions theory extends mental file theory by drawing a distinction between a mental file and its predications, noting coreference requirements on file predications, and applying the apparatus of defaults and defeaters from entitlement epistemology. This supplies the materials for an explanation of referential coordination that meets the requirements and desiderata on an explanation much better than competing explanations. It offers a view of thought that is psychologistic, referentialist, and lightweight, and a take on referential coordination as a matter of coherence that bridges the personal and sub-personal levels. As such, it provides a relatively deflationary view, backed by strong explanatory considerations, of one part of how thought works.

Chapter 7 - Applications

7.1 Introduction

This chapter addresses some matters left open from earlier chapters: the fate of concept theory, the extension of the explanation to identity judgements, and special ways of thinking. The deflationary character of the explanation is consequential on all three topics.

As a lightweight approach to thought, the coordination functions explanation speaks against the existence of concepts as theoretical entities introduced for the purpose of explaining conceptual capacities. This conclusion is complicated by the fact that there are some ways of cashing out what concepts are that are compatible with the coordination functions view. It helps to capture this by adapting the distinction between concepts and *conceptions*. A conception can be understood as a collection of beliefs and other attitudes that a person has concerning something. Conceptions thus contrast with concepts in the non-explanatory sense that captures the capacity to think about something in a coordinated way, and in the explanatory sense of what explains that capacity. Taking conceptions as bodies of referentially coordinated thoughts, it is possible to comprehend much of what goes under the title of the theory of concepts as concerning the nature of conceptions, such as their structure and cognitive role (§7.2).

Referential coordination can be applied to the problems of identity judgements and special ways of thinking by means of conceptions. I argue that coordination functions that are not requirement-based but rather based on information can articulate the empirical significance and rational role of identity judgements. The notion of a *conception link* is introduced to capture this. A conception link imposes a disposition to treat the information contained in two mental files as referring to the same, thus overcoming the fact that they would not otherwise do so. Conception links are introduced by theoretical stipulation, much like mental files themselves, and also made the target of substantive commitments (§7.3). Conceptions also provide a way of comprehending special ways of thinking. On what I call the *deflationary view*, they are not special because of the presence of a special representational mechanism, but and only insofar as they involve conceptions that play a special role. The deflationary view diverges from the more

orthodox view, but deserves consideration due to its being the upshot of the best explanation of referential coordination (§7.4).

7.2 *Concepts and Conceptions*

As I argued in §2.3, concept theory posits entities to explain what makes thoughts that are exercises of the same referential capacity also exercises of the same conceptual capacity. These are capacities that are exercised in thoughts that are referentially coordinated. Because they explain both the referential content of thoughts and the role of thoughts in inference, concept theory posits heavyweight entities. If the coordination functions explanation is correct, then heavyweight theories are incorrect. So we should conclude that concept theory is discredited, and eliminate concepts from our theorising about thought. This provides a useful way to think about the main consequence of the argument for the coordination functions explanation. Many theorists of thought accept that there are concepts, but the coordination functions explanation indicates that this commitment rests on shaky foundations. Some nuance is called for, however.

Concept theory says that there are things that answer the job description of concepts. Concept theories consist of answers to the question (the identification question, as I called it) of what answers to the job description. Answers to the identification question differ, at the most fundamental level, over which side of the vehicle/content distinction concepts are held to fall. Concepts as *contents*, that is, as Fregean senses, serve no purpose on the present showing, and can be eliminated. By contrast, vehicles of thought, mental representations the presence of which account for a proposition being the content of a thought, are not shown to be redundant. Nor is their being typed in systematic ways that reflect the exercise of conceptual capacities. In fact, the coordination functions theory could be used to provide a typology for this purpose. One could call these representations ‘concepts’, though it is not at all essential and, I would suggest, rather misleading: it is not their presence that *explains* why thoughts are exercises of the same conceptual capacity.

Two further clarifications are worth making. The first is that concept theory is to be distinguished from the idea that there are concepts in a non-explanatory or descriptive sense, as when we attribute concepts in order to capture a thinker’s conceptual capacities, the exercises of which are referentially coordinated thoughts. Doing this is not to posit the heavyweight entities of concept theory. These are brought in to explain what makes

for thoughts being exercises of the same conceptual capacity. Moving from talk of concepts in the descriptive sense to concepts in the explanatory sense is easy to do, but is nevertheless a substantive move, and one that should be resisted if my argument here has been correct.

The second is that *psychological* concept theory remains untouched. The relationship between the rejection of philosophical concept theory and psychological concept theory bears more discussion than I can give it here. The immediate point is that although coordination functions theory speaks against philosophical concept theory, it is broadly complimentary with psychological concept theory. As I suggested in §6.2.5 above, coordination functions theory can be made more firmly based in the empirical psychological literature by relating it to the study of information used in categorisation tasks, in the form of prototypes, exemplars, theories, and so on. The extent to which the basic model of mental files might be usefully illuminated is an interesting line of further inquiry that may provide a way of reinforcing the appeal to mental files. It may also have implications for whether the corresponding position of eliminativism about psychological concepts, of the kind advocated by Machery (2009), is defensible.

These two points can be brought out in a way that will prove to be of use in showing how the coordination functions explanation can be addressed to the two desiderata concerning identity judgements and special ways of thinking. Adapting some terminology given psychological currency by Rey (1985; see also his 2010 comment on Machery 2009), we can distinguish between concepts and *conceptions*. Rey's original concern was to criticise what he saw as the mistaken idea that theories of prototypes are theories of concepts, which, roughly, he takes to be the entities posited by philosophical concept theory. By contrast, Rey takes conceptions to be how people 'access' concepts (1985: 260). Whatever this means exactly need not concern us. The distinction as I intend it is between concepts, understood either in the explanatory or descriptive senses, and conceptions as bundles of referentially coordinated information. A conception therefore corresponds to the networks of information linked by requirement-based coordination functions.

Adopting this notion of a conception makes clear how coordination functions theory is complimentary to psychological theories of concepts. Such theories concern the structure and content of conceptions, or of the various forms that conceptions may take when employed in given cognitive tasks. Another benefit is that topics usually taken as apt for

concept-theoretic explanation may instead be transposed into conception-theoretic explanations. That is, one can explain facts about the way that cognitive capacities make use of and depend on different bodies of information, while rejecting any role for heavyweight entities in these explanations. This has particular relevance to the application of coordination functions theory to related problems, including the problems presented by Frege's puzzle about identity and special ways of thinking, as the remainder of this chapter will demonstrate.

On a related topic, the reader may recall the earlier discussion of Fodor's informational atomism (§4.4). This view holds concepts simply to be expressions in a language of thought whose referential contents do not depend on their role in any inference or pattern of categorisation. Many critics of Fodor's theory of concepts have argued (e.g. Peacocke 2000, 2004; Prinz & Clark 2004) that it makes concepts psychologically nugatory; concepts are needed to explain things like rational inference and categorisation behaviour (there is, I note, quite a big difference between these things), and so atomism about concepts must be wrong, or at any rate much worse than competitors that do not have this consequence and so are more explanatorily powerful.

I would like to suggest that the critics are correct, though for the wrong reasons. What Fodor's atomism effectively does, we can now say, is to provide a lightweight theory of concepts in the descriptive sense in terms of a theory of concepts in the explanatory sense. Fodor is right in at least implicitly holding that a theory of concepts in the descriptive sense should be a lightweight theory, because the explanation of referential coordination should be a lightweight one. But his critics are right that one should *not* tie that to a theory of concepts in the explanatory sense. The 'conceptual' part of conceptual atomism is the problem, not the 'atomism' part. Jettisoning the commitment to concepts in the explanatory sense allows us to preserve the lightweight spirit of atomism while avoiding the charge of being psychologically nugatory. The work thought to be done by concepts can be done by conceptions, and the existence of conceptions is perfectly coherent with atomism (cf. Margolis & Laurence 1999: 72-5).

One final point before moving on. One might have the suspicion that the rejection of concept theory being recommended is not wholehearted, because one can make out a concept-theoretic interpretation of the coordination functions interpretation. Such an interpretation would view conceptions understood as the bundles of information linked by coordination functions *plus* whatever it is that makes for intentionality as sufficient for

an explanation of conceptual capacities. In which case, these agglomerations would be concepts. So even if not by itself a theory of concepts, coordination functions theory is not incompatible with such a theory.

The problem is that these agglomerations would not be concepts. As gerrymandered entities, they do not have the kind of explanatory unity one would require of a class of entities to serve as an answer to the identification question; concepts would not be a genuine kind of thing. And identifying these agglomerations as concepts would ignore the point that there is no explanatory relationship between the parts of the agglomeration, something on which traditional theories of concepts place a great deal of importance. In other words, it would ignore the difference between heavy- and lightweight approaches to thought. In fact, this interpretation demonstrates the point of the foregoing remarks about conceptions, that they can do the work that many take to be done by the sort of entities posited by concept theory. The fact that one could give this the somewhat absurd title of a non-concept-theoretic explanation of concepts just goes to show what a tangle it is possible to get into here. The lightweight notion of a conception based on coordination functions provides a way out; I recommend taking it.

7.3 Identity Judgements

Referential coordination is a way for identity to figure in thought in an inferentially basic way. Identity can also figure in a less basic way, as the topic of belief. Such beliefs will reflect a discovery (or apparent discovery) of an identity. One can find out that something one already has thoughts about satisfies some description that can only be satisfied by one thing (as in when one discovers that, say, Cicero is the most famous Roman orator), and one can find out that what one thought was two things are in fact one. Having done so, one can rationally make new inferences. The problem about this less basic way in which identity figures in thought and inference finds expression in the form of Frege's puzzle about identity (see §1.4.2). This puzzle has two parts:

How can identities have empirical significance? Everything is self-identical, and so we know automatically that one thing is identical to itself, and not identical to everything else. But then how is discovering an identity possible, as it evidently is?

How is there a difference in the rational role of identity judgements? Trivial identity judgements are obvious, perhaps knowable *a priori*, whereas informative identity judgements can only be knowable on the basis of empirical discovery. What accounts for this difference?

A good explanation of referential coordination should ideally extend to giving answers to these questions, since the capacity to form identity judgements, at least in the form recognisable in the case of human thought, can be seen as depending on the prior capacity to have referentially coordinated thoughts. This section shows how the coordination functions explanation provides satisfactory answers.

7.3.1 Conception links

Identity judgements can be understood as imposing a kind of coordination function, what I will call *information-based* coordination functions. An identity judgement has the effect, at the information-processing level, of creating a link between conceptions. *Conception links* are like mental files: theoretical entities given a role by stipulation, which are then made subject to substantive commitments. The role they have is to put in place dispositions to make the inferences similar to the paradigmatic form (similar in that they conjoin and generalise, different in that they include an identity premise) across conceptions, rather than within conceptions as in the case of referential coordination proper. These dispositions impose coordination functions with target and base predications belonging to different mental files. These are information-based coordination functions. Information-based coordination functions exploit but are distinct from requirement-based coordination functions, as they do not have a basis in coreference requirements imposed by the dispositions brought about by the presence of a mental file and its predications.

The questions of empirical significance and rational role of identity judgements can then be answered like so:

The empirical significance of identity judgements consists in the establishment of a conception link. Conception links impose information-based coordination functions; these make the coherence of the newly created inferential dispositions with respect to the stock of information a matter of rational sensitivity. Empirically significant judgements create inferential

dispositions not previously had, and so present the possibility of inconsistencies not previously in need of resolution.

The rational role of an identity judgement corresponds to whether it can be such as to impose an information-based coordination functions or not. Trivial identities correspond to cases where all the predications that would be referentially coordinated already belong to a single conception. Informative identities correspond to cases where some predications that would be referentially coordinated do not already belong to a single conception.

Both answers are supported by the same extension of the general idea behind the coordination functions account of referential coordination. Empirical significance, and hence rational role, of identity judgements depends what a thinker's informational set up is like, and not on any representational mechanism. This account therefore concurs with Almog in holding that

informativeness does not rest in the absolute internal informativeness of a proposition we do grasp; it lies in lack of information.... None of this impinges directly on semantics. It is a question in cognitive psychology and human information processing: what information we do have in the head and what it can settle. (2014: xi-xx)

The lack of information is to be understood as the absence of an information-based coordination function drawing together the information that a thinker has. Information can be gained so as to make such a lack incoherent, and in cases where a thinker is rationally sensitive to this, she will form a judgement with the effects just sketched.

The answers to the two questions are not meant to be a radical new departure. The basic idea will be in common to any reasonable account of identity judgements. The basic idea is that empirically significant identities work by bringing together, in some way, what are otherwise distinct ways of thinking, so leading to a change in the overall information state, and difference of rational role consists in whether or not they are apt to be empirically significant in this manner. Empirically significant identities operate on our ways of thinking, and the rational role of identity judgements reflects how this interacts with what inferences we can rationally make. Conception links and information-based coordination functions simply provide a way of substantiating this basic idea. They show

that one can do so without anything like the apparatus of Fregean sense, which might be taken to be mandatory when addressing the role of identity judgements in thought—or at least, to be adopted for want of a better alternative. But a better alternative there is indeed.

There is an obvious similarity between the conception links account of the empirical significance of identity judgements and other treatments of identity judgements in terms of mental files. The idea of links between mental files is a feature of Recanati's discussion of identity judgements within his indexical model of mental files, Recanati refers to the idea, found in Lockwood (1971: 209) and Strawson (1974: 51-6) that identity judgements can be thought of as *merging* mental files, creating one where previously there were two. The difficulties faced by this as an adequate approach are not inconsiderable, not least insofar as it fails to explain how identity judgements can be made with less than absolute certainty, and by extension how they interact with a thinker's with subjective probabilities (Recanati refers us to Lawlor 2001: 62-5 and Millikan 2000: 147-9).

As for Recanati, himself, he holds that there is

an operation on files whose role is precisely to overcome that architectural limitation [i.e. that information associated with distinct files cannot be integrated in inference], by licensing the integration/exploitation of information distributed in distinct files. That operation, following Perry, I call linking. When two files are linked, information can flow freely from one file to the other, so informational integration/exploitation becomes possible. Thus if I learn that Cicero is Tully, this allows me to put together the pieces of information in the two files..

To accept the identity 'A=B' is to link the two files corresponding to the terms on each side of the equals sign. It would be incoherent to accept the identity 'Cicero = Tully', and not let the information in the respective files get together and breed. (2012: 43-44)

Putting aside for a moment whether Recanati-Perry linking is the right way to think about the empirical significance (and also that Recanati does not explain why linking allows one to perform the inference, compare §5.5), it must be seen as distinct from conception linking. For one thing, nothing about conception links implies that

information can flow freely. That is, for any predication of a file P that is in a conception C , if there is a conception link between C and a distinct conception C^* , there will not necessarily be a predication P^* in C^* that carries the same information as P . A deeper difference is that, whereas Recanati-Perry linking is what we might call a file operation (much like the expansion, detachment, and so on discussed in §5.4), changing as it does the functional profile of mental files, conception links make no such changes to mental files, but rather to the functional properties of their predications. So although there is some similarity, Recanati-Perry linking and conception links to create an information-based coordination function are quite different.

Recanati does not elaborate much further on the topic of linking, other than saying that it is “a quite fundamental operation... involved, for example, in the phenomenon of recognition (which involves linking a perceptual file and a file based on memory)—or at least, in some forms of recognition.” (op. cit.: 44.) Recanati-Perry links can be treated as causal primitives, much as mental files and their predications are. This goes for conception links as well. We can say that conception links are whatever they need to be so as to create dispositions of the paradigmatic kind across conceptions (on Recanati’s view these links are between files rather than conceptions, though they can be treated as more or less the same thing). It might be objected that this is to leave things much too underspecified. Even though it is legitimate to bring them in as primitives, just as mental files are, they need to be made the target of substantive claims that allow them to be open to assessment as part of our overall best theory of the mind.

It might seem unclear how this could be done. I suggest a focus on the point made about merging made above, that it fails to take into account the fact that identity judgements can be made with less than absolute certainty. This point can be captured in terms of conception links, like so:

Strength: The strength of the conception link between two conceptions C and C^* tends to reflect the subjective probability function P (>0) assigned to the proposition that C and C^* are conceptions of the same thing.

The strength of a conception link is how likely it is that a thinker will revise it in the face of new evidence. The subjective probability function P will, as usual, deliver values between 0 and 1. One conception link is stronger than another if the value of P is higher, weaker if the value of P is lower; hence the strength of the conception link reflects P .

Values of P less than 1 correspond to the subjective probability that there are inconsistencies between the information in C and the information in C^* , modified by the probability that the contradiction can be resolved consistent with the other information. A value of 1 will therefore be effectively the same as merging, and represents the extreme case of it being excluded that a thinker's information indicates that there are two things rather than one. In such a case, there is a more profound shift in the informational state, wherein two conceptions are fused into one. There will be no rational way back to the previous information state without reconstruction, in the sense introduced in §6.2.4. Whether a conception link with strength reflecting $P=1$ is ever obtained in actuality (or for that matter, whether there are uninformative identity judgements, about which some scepticism is warranted; what would be the point in them?) can remain a moot point.

Strength can be extended to the claim that coordination functions based on information indicating identity interact with a thinker's broader stock of information in line with the Bayesian calculus. The extent to which this is so will be modified by the heuristics and rules of thumbs employed in updating this information. This is why Strength says that conception links will *tend* to reflect P , rather than reflect it without qualification. There is no point pretending that there will be any simple account of the exact nature of the cognitive links in question. But the basic idea, that identity judgements having empirical content can be cashed out with an application of the apparatus of subjective probability, finds a neat fit with coordination functions theory. Touching again on a moral drawn at the end of the last section, the coordination functions account of the empirical significance and rational role of identity judgements allows for a high degree of complexity. The nature of conception links will be at least as complex as global probabilistic information update, which is very complex indeed.

7.3.2 Advantages

If we assume that a mental file approach to identity judgements must be on the right lines, then which model should we prefer, conception linking or Recanati-Perry linking? An advantage of conception links as characterised by Strength is that they offer a more realistic picture of how explicit judgements of identity work, both in how they interact with both requirement-based coordination functions, and so with the referential coordination of thoughts, and in how facts about identity can be discovered and integrated into a thinker's rational conception of things.

One way in which this shows up is in what we might call the *pondering problem*. This problem is nicely brought out by Papineau and Shea's critical notice of Millikan (2000):

We have attitudes to propositions of identity other than acceptance or rejection. We can entertain such a proposition, and wonder about it. (*Is Cary Grant the same person as Archie Leach?*) [...] Am I quite sure that Jim was the man I saw robbing the grocery store? Maybe water isn't H₂O after all. (Papineau & Shea 2002: 463-4)

The problem is to allow for the fact that one can also entertain identity judgements, presuppose them for the sake of argument, ponder them, and so on. The pondering problem is therefore to account for the rational role of *merely entertained* identity judgements, and constitutes an extension of the rational role question.

Conception links can handle the pondering problem neatly, by bringing in the idea that conception links can be simulated. Such simulations simply involve computing the subjective probability assigned to the proposition that the information in two conceptions is information about the same thing, without actually changing that subjective probability. Entertaining an identity judgement is computing these changes. By contrast, file operations such as Recanati-Perry linking leave this aspect of the rational role of identity judgements untouched without an adaptation, but also potentially render them problematic.

Suppose we apply the same simulation idea, which is the most obvious adaptation. Instead of computing changes to subjective probabilities about information, a file operation account would need simulations of wholesale changes to the whole mental file system. If two files are linked so that information can flow freely, then this could potentially lead to a wholesale change in a thinker's information. Simulating these changes would need to include assessing all such changes for plausibility. This would require a potentially massive computational effort (compare Lawlor 2001: 56-7). This is another instance of the problem presented by global inference. Merely entertained identity judgements would be far more drastic and costly than they need to be. In the first instance, only one proposition should be entertained, not potentially the whole of a thinker's store of information. Simulated conception links do not have this consequence,

and Recanati-Perry linking does. Conception links should therefore be preferred over Recanati-Perry linking.

A related difficulty faced by heavyweight explanations more broadly concerns how the explanatory entities they employ should be individuated so as to allow for a smooth explanation of referential coordination before and after information concerning identities has been gathered. Before, there are two such entities, but after making a judgement that the two things are in fact one, the thinker is now disposed to treat her thoughts previously treated as being about different things as being about the same thing. What then happens to the two entities? Do they remain distinct? If so, then it makes the prediction that the thinker is not disposed to treat the thoughts as being about the same thing, which is false to the rational role of identity judgements. If not, then it makes the false prediction that the information concerning identity becomes trivial, which is false to their empirical significance.

The import of this problem is shown in some difficulties raised for Recanati's indexical model (see Goodsell 2013; Onofri 2015), to which he has responded (Recanati 2013) by accepting that the model needs additional elements, which he calls 'referential indices', complicating an already elaborate ecosystem of mental files to account for change in ways of thinking. Sense theory has related difficulties (cf. Kripke, 1979; Speaks 2014), and parallel problems will arise for mental representation typologies (recall the problems faced by the neo-empiricist typology regarding inferences involving change over time; see Millikan 2011 for related problems addressed to the originalist typology).

The point is not that these problems are insurmountable, but rather that heavyweight theories seem to require additional resources to provide an account of the rational role of identity judgements (the fact they are often addressed to identity judgements in the first instance makes this rather awkward). The cornerstone of a lightweight approach is that it provides a basis for referential coordination that is not tied to the epistemology and psychology of identity judgements. This allows these latter problems to be subsumed under models of probabilistic global informational update, and implementation proposals subject to requirements of descriptive adequacy and psychological realism (perhaps involving heuristics, biases, mental models, and so on, of the sort discussed by Nickerson 2004, and recently in more popular format by Kahneman 2011). This is perhaps counter-intuitive, but it means that the coordination functions explanation

needs no additional resources to account for how discovering identities can interact with our attempts to gain reliable information.

The fact, considered above as a potential problem, that lightweight explanations do not make patterns of referential coordination transitive helps rather than hinders in this regard. It also means that the epistemology of identity and the actual psychological processes that implement it can be held apart, as neo-Fregeans motivated by anti-psychologistic inclinations generally insist, while at the same time avoiding the problem that comes with a blanket anti-psychologistic approach that leaves it obscure how they can be related.

7.3.3 Summary

The coordination functions explanation of referential coordination can be extended to identity judgements in a way much like the introduction of mental files: conception links are introduced by theoretical stipulation as what impose inferential dispositions, that ground a particular form of coordination function; they are then made the topic of substantive commitments, in this case, regarding how they interact with a thinker's subjective probabilities. The basis of this account in a lightweight view of thought gives it some advantages over contrasting accounts. I conclude that the coordination functions explanation meets the fourth desideratum.

7.4 *Special Ways of Thinking*

The idea of a conception can be applied to the problem of special ways of thinking. More accurately, it can be used to show that there is not so much of a problem there at all. I propose a deflationary view of special ways of thinking, on which they involve conceptions that are special only in that they play a special role. There are reasons to prefer this to the more orthodox view of special ways of thinking, one that sees them as tied to special contents or representational mechanisms. The deflationary view might be seen as unsatisfactory by those who hold that demonstrative thoughts in particular have an indispensable role to play in our having thoughts about the world at all, but I argue that the extent to which there is a problem for the deflationary view on that front is minimal.

7.4.1 The deflationary view

What I earlier called special ways of thinking are the rational relations between thoughts that are employed in certain cognitive capacities. These thoughts concern things or properties other than oneself in one's immediate environment that figure in one's perceptual experiences, and first personal or *de se* thoughts concerning oneself. Such thoughts are often categorised as demonstrative or indexical, and are often understood in terms of indexical and demonstrative concepts. The precise characterisation of these thoughts and/or concepts, and how or whether they relate to each other, is a controversial matter. What is generally not taken to be controversial is that there are such things. And it is usually supposed that there is a close link between them and particular classes of referring expressions, such as indexical and demonstrative referring terms, adverbial expressions, and inflections such as grammatical person and tense.

These two claims—that there are these classes of thoughts or concepts, and that understanding them goes hand in hand with understanding the semantics of these classes of referring expressions—go together to make up what we might call the *semantic view* of special ways of thinking.

I propose a conception-based alternative to the semantic view. On this alternative, special ways of thinking can be understood by employing the idea of a conception like so:

The deflationary view: Thoughts exhibiting special ways of thinking are ordinarily referentially coordinated thoughts (i.e. involve file predications that are target-base pairs for coordination functions) that belong to conceptions used for special purposes, for example, recognition of perceptually salient objects, action guidance, the sustenance of a rich self-concept, patterns of emotional response, and so on.

I call this the deflationary view because it finds the specialness of the thoughts that exhibit special ways of thinking only in the role played by information that makes up a single conception; the fact that these thoughts exhibit sameness in ways of thinking, that it is not enough to account for the rationality of inferences involving them to simply advert to their referential content, is to be explained in the same way as any other group of thoughts that exhibit sameness in ways of thinking.

The deflationary view therefore contrasts with the semantic view because it denies any role to special representational mechanisms, and does not link special ways of thinking with anything to do with the semantics of indexicals, demonstratives, and the rest. What is special about special ways of thinking is that the bundles of referentially coordinated thoughts play special roles, not what makes them bundles. This compares with Cappelen and Dever's (2013: 58) proposal to see the alleged phenomenon of essential indexicality as instead the more general feature of referential opacity or non-substitutability, and Millikan's claim that "what makes [indexical thoughts] special is not their semantics but their functions, their psychological roles, their impacts on my behaviour." (1990: 733)

As rational creatures, with projects and values that motivate action that requires sensitivity to our surroundings in the broadest sense, these uses are special because essential to our lives. The ways of categorising thoughts on the semantic view correspond to these basic ways of interacting with things. So-called demonstrative thoughts concern information gained and stored in short-term memory for purposes of recognition and guidance of immediate action. First personal thoughts concern information about oneself needed to guide action in both the short- and long-term, information that interacts with one's intentions and practical abilities to make certain courses of action rational. First person thoughts are clearly essential to the idea of a self-concept, taken from social psychology, consisting not just of beliefs about one's characteristics, but beliefs concerning one's place in broader social narratives, and one's self-worth and competences in relation to those narratives, within which one can make sense of one's actions (cf. Oyserman, Elmore & Smith 2011). Emotional response, and particularly self-conscious emotional response (cf. O'Brien 2011), is not obviously reducible to action or belief about the self, but is no less a part of how we respond to our surroundings. Though perception and action are often taken to be central in the philosophical literature, it is hard to imagine a realistic picture of an actual human life lacking social narrative and emotional response, and they deserve their place in the account of special ways of thinking as much as the more mundane tasks of self-location.

7.4.2 Against the semantic view

Are there reasons to prefer the deflationary view over the semantic view? It helps first to say a little more about what the contrast involves.

Seen as an approach to the psychology involved in special ways of thinking, the semantic view models thoughts with these roles on the meaning and use of indexical and demonstrative expressions and other syntactic forms. Not surprisingly, proponents of the semantic view divide along lines familiar from debates about meaning. One wing is represented by those, taking their cue from Frege (1918/1956), who seek to extend the notion of sense in some form. This extension has often been developed in response to critics from the other wing, which, following Kaplan's (1989) seminal treatment of indexical and demonstrative expressions, finds expression under the slogan of 'direct reference'. The direct reference wing makes no appeal to special propositional contents. Instead, it posits special non-propositional ways of determining truth-evaluable items of a non-Fregean sort (usually Russellian propositions, composed of objects, properties, and relations themselves rather than combinations of Fregean senses, though propositions as sets of possible worlds also feature). Along with Kaplan's (op. cit.) character/content distinction, the direct reference wing of the semantic view has been developed by, for example, Lewis' (1979) proposal to see attitudes as self-ascriptions of properties regarding ones location in 'logical space', and Perry's arguments against the neo-Fregean account of demonstratives (1977).

As for the neo-Fregean rejoinder, Perry's arguments against Frege's treatment of demonstratives was the target of Evans' (1980) vigorous defence, in which he introduced the idea of demonstratives as expressing 'dynamic modes of presentation', subsequently developed in his (1982). Evans suggests that the core of the Fregean notion of sense can be preserved, with some modifications: dynamic modes of presentation are said to depend on a thinker's perceptual tracking of objects through time and space (on which more presently). Despite their differences, when applied to the case of thoughts, the neo-Fregean and direct reference wings are united in the idea that there are reproducible 'templates' of how information is processed, and these templates serve to both determine a reference and explain facts about the rationality of information processing. Either these templates are held to serve as something akin to Fregean propositions, or they are not, but the basic idea remains in place either way.

Aside from Boer and Lycan's early expression of scepticism about the specialness of first person thoughts ("attitudes *de se* are simply attitudes *de* their owners." 1980: 432), opposition to the semantic view has been confined to the more recent (as of 2016) literature. The general tenor of these interventions is primarily negative, mainly concerned to deny that there is any particular phenomenon worthy of attention. This

typically takes the form of taking issue with Perry's claim concerning 'the essential indexical' (Tiffany 2000; Millikan 2012, somewhat recapitulating her 1990; Magidor forthcoming). A radical example of such opposition is in Cappelen and Devers' book length treatment (title: *The Inessential Indexical*), the purpose of which, they write, is to

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. (2013: 3)

The deflationary view as I have outlined it is obviously of a piece with this, though it comes with some distinctive credentials; it is the most simple extension of the coordination functions explanation, which is the best theory going of how our thoughts are referentially coordinated. It is thus founded on a positive proposal, rather than simply a negative reaction to the semantic view.

How exactly the deflationary view is formulated is a matter for terminological decision. On the one hand, the point of categorising thoughts as indexical or demonstrative, or first-person or what have you, might be just to provide a way of picking out the roles played by special conceptions. The point is that, on the deflationary view, the categorisations are of no significance when it comes to explaining what makes those thoughts belong to those conceptions, nor of what gives those conceptions their role. For the purposes of processing perceptual information and forming intentions, *some* conception has to be employed, and the question of why a particular conception plays that role is misplaced. If, on the other hand, the point of categorising some thoughts as indexical, demonstrative, or first-person is to mark that they *are* of significance in this way, then the fact that there is no need for a special explanation of how thoughts get referentially coordinated can be expressed as the claim that there are no demonstrative or indexical thoughts.

Weight of philosophical opinion is on the side of the semantic view. There is, however, at least one good reason for thinking that the semantic view is likely false, independent of considerations to do with the best account of referential coordination.

One of the distinctive functions of indexical and demonstrative expressions is to allow for the fact that utterances and inscriptions are often made to express to an audience how someone, the speaker or inscriber, sees things, both literally and figuratively. These expressions help to coordinate different points of view, or perspectives:

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*. (Cappelen & Dever 2013: 1)

In these cases, for example, the grammatical first person provides a simple and quick way for reference to the speaker to be effected, in a way that supplies any potential audience with an easy knowledge of what, or who, the subject of the sentence is, without presupposing their possession of any further information about the speaker (such as her name). Similarly, demonstrative expressions provide a simple and quick way for reference to salient items in the environment to be effected, against a shared background of assumptions about what, in the circumstances, would be salient. The fact that these assumptions can be highly complex and open ended makes the game of coming up with proposals about what this salience involves an interesting one, though I suspect ultimately fruitless (cf. Mount 2008).

When it comes to thought, however, the question is why anything like these semantic mechanisms for coordinating different points of view should be needed. Thoughts occupy a single point of view. They are thought by the thinker with that point of view. One does not need to coordinate different points of view because one is not *communicating* anything in having a thought. The same goes for thoughts that are involved in perception and action. Thought represents a distinct engineering problem, one that has no call for semantic shortcuts. What work would analogues of the semantics of these expressions need to do?

That there is a need of this kind is usually simply assumed by proponents of the semantic view. As Cappelen and Dever (2013) suggest, actual arguments for this tend to be thin on the ground. Perhaps it is based on the assumption that thought and language are merely different vehicles for what must be the same contents, together with the view that indexicals and demonstratives express special contents. This is unavailable if we go with

the direct reference wing of the semantic view, and we have seen reasons to suspect that the neo-Fregean treatment of thought is explanatorily lacking. There is perhaps also a pervasive idea that thought is something like one's internal monologue, in which we internally utter words, including words like 'I' and 'here', and these internal utterances are as much in need of interpretation, and so call for the same semantic analysis, as ordinary utterances made out loud. But internal monologue is imagined speech, imagined expression of thought. Thoughts do not stand in need of interpretation by the thinker or anyone else. They only need to be *thought*. Semantic shortcuts don't get a look in.

7.4.3 An inflationary objection

Many proponents of the semantic view do not just hold that there are indexical, demonstrative, or first person thoughts as a matter of psychological fact, but rather take their existence to be crucial for our cognitive lives. This may seem to provide a decisive objection to the deflationary view, to the effect that a more inflationary account is needed for them to do this. I want to suggest in response that this can be made consistent with the deflationary view, so long as one allows that the work can be done by the conceptions that play the special roles.

Demonstrative thoughts, or more exactly, demonstrative concepts employed in those thoughts, are the most clear case of the more inflationary role given to thoughts exhibiting special ways of thinking. Proponents of the indispensability of demonstrative concepts argue that, just as we can say something about an object in virtue of employing a semantic mechanism that makes some salient thing the subject of an utterance, so we can think about an object in virtue of employing a representational mechanism that makes some salient thing the subject of a thought. That we can do this explains how we can have a rational conception *of* things at all. Without such thoughts, our rational conception would make no meaningful contact with the world, populated as it is by particulars located and related to each other in space and time. Perhaps all we could have otherwise are descriptive thoughts, employing descriptions which particular objects may or may not satisfy. Perhaps we could not have even that. A theory of what the representational mechanism is, and how exactly it serves to fix reference, is held to answer to this explanatory aspect of demonstrative thoughts.

The classic expression of the idea that demonstrative thoughts, or more exactly so-called *perceptual* demonstrative thoughts where perception is what makes objects salient or ‘available to thought’, are crucial to understanding how we have a rational conception of the world as populated by particulars located in space and time can be found in Strawson’s *Individuals*:

We think of the world as containing particular things some of which are independent of ourselves; we think of the world’s history as made up of particular episodes in which we may or may not have a part; and we think of these particular things and events as included in the topics of our common discourse, as things about which we can talk to each other... (1959: 15)

[T]here are many cases of identification falling under this condition. An expression is used which, given the setting and accompaniments of its use, can properly, or at least naturally, be taken, as then used, to apply only to a certain single member of the range of particulars which the hearer is able, or a moment before was able, sensibly to discriminate, and to nothing outside that range. Cases of this kind are the cases, par excellence, for the use of demonstratives.... I shall say, when this first condition for identification is satisfied, that the hearer is able directly to locate the particular referred to. We may also speak of these cases as cases of the *demonstrative identification* of particulars. (op. cit.: 18)

Evans (1982: 141-191) stresses the need to go beyond Strawson’s account to make explicit the rational dimension of perceptual demonstrative thoughts. Campbell (2002) has argued that demonstrative thought specifically requires sensory attention; Dickie (2011) and Smithies (2011) extend this to the idea that demonstrative thought needs to be understood in terms of the epistemic role of perception. Related views can be found in Peacocke (1981; 2008: 48-51) and McDowell (1990; 1994).

The idea in common across these advocates of the importance of demonstrative thoughts is that they supply the *knowledge of reference* necessary for reference to particulars. Knowledge of reference is held to require knowledge of what individuates the particular in question; that is, to know the condition sufficient for making it different to all other things. A thing’s location in space and time individuates it, and perceptual demonstrative

thought provides the contact with things needed to have such knowledge (this is what, as I understand it, is meant by the expression ‘perceptually-based thought’, frequently used to gloss the idea of perceptual demonstratives). Such knowledge therefore bears on the identity of what is thought about, and so can explain for instance, why we might think something about a perceived object, yet not know that that judgement concerns something we had only been told about. Or, conversely, it can explain how one can put together information about a perceived object and rationally judge that it is the same as something one has heard about. And it can explain how we can rationally treat multiple thoughts as being about the same thing, when one and the same object is made salient through several perceptual episodes. The claim that demonstrative thoughts are indispensable therefore has a distinctly neo-Fregean character, and indeed the proponents of this view named above often express it in terms of Frege’s notion of sense.

Granting the claim that, if we did not have demonstrative thoughts, our having a rational conception of things that makes genuine contact with how things are would be impossible, the result seems to be that special ways of thinking are special in a way that the deflationary view cannot encompass, as it involves nothing that could supply knowledge of reference of the sort held to be provided by demonstrative concepts.

Should we grant the claim? It is not at all clear that there is a way to critically assess something so profound on its own terms (the fact that arguments for the existence of demonstrative thoughts tend to be thin on the ground makes things more difficult on this score). But it is possible to assess it by bringing it into slightly more familiar territory, by dint of assessing the more general approach to ways of thinking that the more inflationary role is held to entail. We should ask whether it is really the case that, were identity relations concerning what we think about separated from facts about perceptual salience, then we would have no way of accounting for the rationality of our conception of things.

The answer we are now in a position to return is a simple no. In fact, coordination functions theory is an adequate theory of referential coordination, and this is wholly independent of anything along the lines of perceptual demonstration. The conditions sufficient for thoughts being referentially coordinated make no reference to perceptual experiences. ‘Perceptually-based thought’, in anything stronger than the fairly banal sense of ‘thought involving information that was in fact gained through perception’, is a category without work to do.

This may seem simply incredible to proponents of the semantic view for whom demonstrative thought represents the meeting point of mind and world. Again, the disagreement is deep enough that the prospect of settling it by seeing out the consequences of mutually agreed-upon premises is dim. This does not mean the issue cannot be settled at all. So long as it is agreed that the issue is one of better or worse explanations, this must tell somewhere. The idea is that it tells most keenly where things turn on the explanation of referential coordination. On that score, the tradition that sees demonstrative concepts as playing a crucial role comes off worst, due to its adherence to the semantic view.

We should also ask whether this inflationary objection really presents a problem for the deflationary view. The inflationary objection presupposes that the crucial knowledge of reference is supplied by perceptual contact with things in the world. This only presents a problem if such knowledge cannot be accommodated by the deflationary view, which is not clearly the case. The deflationary view can accommodate this, albeit as part of a piecemeal approach to the rational, epistemic, and intentional properties of thoughts that exhibit special ways of thinking, rather than the unified approach premised on the neo-Fregean view of mental content.

The deflationary view is consistent with the claim that the semantic content of some, perhaps even most, of our thoughts depend on information gained through perception, in such a way that they could not be had were the thinker not to have undergone some perceptual experience. And, more importantly, it is consistent with the claim that there are instances where perceptual experience provides information that makes for knowledge of reference. The obvious suggestion is that it is in the *content* of *conceptions* that such knowledge should be located. Rather than demonstrative concepts, one has otherwise ordinary conceptions put to special purposes. These special purposes can include their providing the materials for knowledge of reference. No commitment to demonstrative concepts is thereby incurred.

A more satisfactory treatment of this topic would need to address in significantly greater detail how exactly to interpret the claim that thought requires knowledge of reference supplied by perceptual experience, or indeed whether there is any viable interpretation of it to be had (cf. Hawthorne & Manley 2012: 71-92). Providing answers to either question lies well beyond the scope of the present discussion. For the time being, we can

conclude that there is no reason to think that the inflationary objection points to any clear problem with the deflationary view.

7.4.4 Summary

The coordination functions explanation supports a deflationary view of special ways of thinking, on which they are special only in that thoughts that exhibit them make up conceptions that play a special role in our lives. This view of indexical and demonstrative thoughts by itself offers no explanation of the capacity to think about the world, as some have taken demonstrative thoughts to do. This may be taken as a reason to reject the deflationary view, but the notion of a conception may be used to do much of that work, and so such a reaction would be unwarranted. I conclude that the coordination functions explanation meets the fifth desideratum.

7.5 Conclusion

The coordination functions explanation of referential coordination undermines the commitments of philosophical concept theory, although the distinction between concepts and conceptions provides a way of preserving some of what goes under that heading within the scope of a lightweight approach to thought. Conceptions can be linked, so allowing the explanation to be extended to identity judgements. Conceptions also provide the materials for a deflationary view of special ways of thinking. On all three topics, there is a trade-off to be made between the relative adequacy of an explanation of referential coordination and the extent of the theoretical ambitions that can be supported by it. Particularly in the case of special ways of thinking, this trade off may seem rather severe. But it is no less attractive once one considers the problems with alternatives.

Conclusion

This essay has explored how we should think about referential coordination. Chapter one set out the problem as one of giving an account of an inferentially basic though personal-level indication of sameness of reference, sufficient for making rational the disposition to engage in a paradigmatic form of inference. Chapter two related this problem to debates in the theory of concepts. Chapters three, four, and five evaluated some prominent theories of thought in relation to the problem, and found them wanting. Chapter six presented the coordination functions explanation, and showed how it does better than the competitors. Chapter seven showed how coordination functions theory can be applied to some closely related problems.

The conclusion in favour of the coordination functions explanation is contingently correct if correct at all. The argument given does not pretend to be conclusive. Explanations can always be improved upon, and that goes for the one recommended here no less. Even so, I submit that the case is a compelling one.

Bibliography

- Almog, J. 2014. *Referential Mechanics: Direct Reference and the Foundations of Semantics*. Oxford: OUP
- Ayedede, M. 1998. 'Fodor on Concepts and Frege Puzzles'. *Pacific Philosophical Quarterly* 79(4): 289-294
- Bach, K. 1987. *Thought and Reference*. Oxford: Clarendon Press
- Ball, D. 2015. 'Indexicality, Transparency, and Mental Files'. *Inquiry* 58(4): 353-367
- Barsalou L. 1999. 'Perceptual Symbol Systems'. *Behavioral and Brain Sciences*, 22: 577-660
- Beck, J. 2012. 'The Generality Constraint and the Structure of Thought'. *Mind* 121(483): 563-600
- Boer, S. & Lycan, W. 1980. 'Who, Me?' *The Philosophical Review* 89(3): 427-466
- Brentano, F. 1874/1973. *Psychology from an Empirical Standpoint*. A.C. Rancurello, D.B. Terrell, & L. McAlister (trans.) London: Routledge
- Brock, S. & Mares, E. 2014. *Realism and Anti-Realism*. London: Routledge
- Burge, T. 1993a 'Concepts, Definitions, and Meaning'. *Metaphilosophy* 24(4): 309-325
- 1993b. 'Content Preservation'. *Philosophical Review* 102(4): 457-488
- 2004. *Truth, Thought, and Reason: Essays on Frege*. Oxford: OUP
- 2013. 'Gottlob Frege: Some Forms of Influence'. in M. Beaney (ed.) *The Oxford Handbook of the History of Analytic Philosophy*. Oxford: OUP: 355-383
- Campbell, J. 1987/88. 'Is Sense Transparent?' *Proceedings of the Aristotelian Society* 88: 273-292
- 1994. *Past, Space, and Self*. Cambridge, MA.: MIT Press
- 2002. *Reference and Consciousness*. Oxford: OUP
- Cappelen, H. & Dever, J. 2013 *The Inessential Indexical*. Oxford: OUP
- Carey, S. 1985. *Conceptual Change in Childhood*. Cambridge, MA.: MIT Press
- Carnap, R. 1947. *Meaning and Necessity*. Chicago: University of Chicago Press
- Casullo, A. 2009. 'What is Entitlement?' *Acta Analytica* 22(4): 267-279
- Chalmers, D. 2002. 'On Sense and Intension'. *Philosophical Perspectives* 16 Language and Mind: 135-182
- 2004. 'Epistemic Two-Dimensional Semantics'. *Philosophical Studies* 118(1-2): 153-226
- 2006. 'The Foundations of Two-Dimensional Semantics', in *Two-Dimensional Semantics: Foundations and Applications*, M. Garcia-Carpintero & J. Macia (eds.). Oxford: Oxford University Press

- 2011. 'Propositions and Attitude Ascriptions: A Fregean Account'. *Noûs* 35(4): 594-639
- 2012. *Constructing the World*. Oxford: OUP
- Dennett, D.C. 1969. *Content and Consciousness*. London: Routledge
- 1987. *The Intentional Stance*. Cambridge, MA.: MIT Press
- 1991. 'Mother Nature vs. the Walking Encyclopaedia: A Western Drama', in W. Ramsey, S. Stich, and D. E. Rumelhart, (eds.), *Philosophy and Connectionist Theory*. Hillsdale, NJ: Erlbaum
- Devitt, M. 1989. 'Against Direct Reference'. *Midwest Studies in Philosophy* 14(1): 206-240
- Dickie, I. 2010. 'We Are Acquainted With Ordinary Things', in R. Jeshion (ed.), *New Essays on Singular Thought*. Oxford: OUP: 213-245
- 2011. 'Attention Fixes Reference By Eliminating Referential Luck', in C. Mole, D. Smithies & W. Wu (eds.), *Attention: Philosophical and Psychological Essays*. Oxford University Press: 292-322
- Donnellan, K. 1970. 'Proper Names and Identifying Descriptions'. *Synthese* 21(3/4): 335-358
- 1974. 'Speaking of Nothing'. *Philosophical Review* 83(1): 3-31
- Dretske, F. 1981. *Knowledge and the Flow of Information*. Cambridge, MA.: MIT/Bradford Press
- 2000. 'Entitlement: Epistemic Rights Without Epistemic Duties?' *Philosophy and Phenomenological Research* 60(3): 591-606
- Dummett, M. 1959. 'Truth'. *Proceedings of the Aristotelian Society* 59(1):141-162
- 1975. 'What is a Theory of Meaning?' in S. Guttenplan (ed.), *Mind and Language*. Oxford: Oxford University Press
- 1978. 'Can Analytical Philosophy be Systematic, and Ought it to Be?' in *Truth and Other Enigmas*, London: Duckworth
- 1991. 'The Relative Priority of Thought and Language', in *Truth and Other Enigmas*. Oxford: OUP
- Edwards, K. 2009. 'What Concepts Do'. *Synthese* 170: 289-310
- 2010. 'Unity amidst heterogeneity in the theory of concepts'. *Behavioural and Brain Sciences* 33(2/3): 210-211
- Egan, F. 1999. 'In Defence of Narrow Mindedness'. *Mind and Language* 14(2): 177-194
- Evans, G. 1982. *Varieties of Reference*. Oxford: OUP
- 1985. 'Understanding Demonstratives', in *Collected Papers*: 291-321
- Field, H. 1978. 'Mental Representation.' *Erkenntnis* 13: 9-61

- Fine, K. 2007. *Semantic Relationism*. Oxford: Blackwell
- Fodor, J. 1987. *Psychosemantics*. Cambridge, MA.: MIT Press
- 1990. *A Theory of Content and Other Essays*. Cambridge, MA.: MIT Press
- 1998. *Concepts: Where Cognitive Science Went Wrong*. Oxford: OUP
- 2008. *LOT 2: The Language of Thought Revisited*. Oxford: OUP
- Forbes, G. 1989. 'Cognitive Architecture and the Semantics of Belief'. *Midwest Studies in Philosophy* 14(1): 84-100
- 1990. 'The Indispensibility of Sinn'. *Philosophical Review* 99(4): 535-563
- Frege, G. 1892/1948. 'Sense and Reference'. *The Philosophical Review* 57(3): 209-230
- 1918/1956. 'The Thought: A Logical Inquiry'. *Mind* 65(259): 289-311
- Friend, S. 2011. 'The Great Beetle Debate: A Study in Imagining with Names'. *Philosophical Studies*, 153(2): 183-211
- Goodsell, T. 2013. 'Mental Files and their Identity Conditions'. *Disputatio* 5(36): 177-190
- Grice, P. 1969 'Vacuous Names', in D. Davidson and J. Hintikka (eds.), *Words and Objections*, Dordrecht: Reidel
- Hall, J.K. 2013. 'Acquaintance and Mental Files'. *Disputatio* 5(36): 119-132
- Hawthorne, J. & Manley, D. 2012. *The Reference Book*. Oxford: OUP
- Heck, R. 2012. 'Solving Frege's Puzzle'. *Journal of Philosophy* 109(1/2): 132-174
- Heim, I. 1983. 'File Change Semantics and the Familiarity Theory of Indefinites'. in R. Buerle, C. Schwarze, and A.v. Stechow (eds.), *Meaning, Use, and Interpretation of Language*. Berlin: de Gruyter: 164-189
- Hornsby, J. 2000. 'Personal and Sub-personal: A Defence of Dennett's Early Distinction'. *Philosophical Explorations* 3 (1): 6-24
- Ichikawa, J.J., & Jarvis, B. 2013. *The Rules of Thought*. Oxford: OUP
- Jeshion, R. 2009. 'The Significance of Names'. *Mind and Language* 24(4): 370-403
- Jeshion, R. (ed.) 2010. *New Essays on Singular Thought*. Oxford: OUP
- Kahneman, D. 2011. *Thinking, Fast and Slow*. New York, NY: Farrar, Straus, and Giroux
- Kahneman, D., Treisman, A., & Gibbs, B. 1992. 'The Reviewing of Object Files: Object-Specific Integration of Information'. *Cognitive Psychology* 24: 175-219
- Kaplan, D. 1989. 'Demonstratives', in J. Almog, J. Perry & H. Wettstein (eds.), *Themes from Kaplan*. Oxford: OUP
- 1990. 'Words'. *Aristotelian Society Supplementary Volume* 64: 93-119
- Karttunen, L. 1976. 'Discourse Referents', in J. D. McCawley (ed.), *Syntax and Semantics* 7: 363-386

- King, P. 2007. 'Rethinking Representation in the Middle Ages: A Vade-Mecum to Mediaeval Theories of Mental Representation', in H. Lagerlund (ed.), *Representation and Objects of Thought in Medieval Philosophy*. Aldershot: Ashgate
- Kripke, S. 1979. 'A Puzzle About Belief', in A. Margalit (ed.), *Meaning and Use*. Dordrecht: Reidel: 239-283
- 1980. *Naming and Necessity*. Cambridge, MA: Harvard University Press
- Lawlor, K. 2001. *New Thoughts About Old Things: Cognitive Policies as the Ground of Singular Concepts*. London: Routledge
- 2005. 'Confused Thought and Modes of Presentation'. *Philosophical Quarterly* 55(218): 21-37
- 2010. 'Varieties of Coreference'. *Philosophy and Phenomenological Research* 81(2): 485-496
- Lewis, D. 1979. 'Attitudes De Se'. *Philosophical Review* 88(4): 513-543
- Lockwood, M. 1971. 'Identity and Reference', in M. Munitz (ed.), *Identity and Individuation*, New York: New York University Press: 199-211
- Loewer, B. 1997. 'A Guide to Naturalizing Semantics', in: B. Hale & C. Wright (eds.) *A Companion to the Philosophy of Language*. Oxford: Blackwell: 108-126
- Machery, E. 2009. *Doing Without Concepts*. Oxford: OUP
- Magidor, O. forthcoming. 'The Myth of the De Se'. *Philosophical Perspectives*
- Margolis, E. & Laurence, S. 1999. 'Concepts and Cognitive Science', in E. Margolis & S. Laurence (eds.), *Concepts: Core Readings*. Cambridge MA: MIT Press
- 2003. 'Concepts'. In S. Stich & T. Warfield (eds.), *Blackwell Guide to Philosophy of Mind*. Oxford: Blackwell. 190-213
- 2007. 'The Ontology of Concepts: Mental Representations or Abstracta?' *Noûs* 41(4): 561-593
- 2010. 'Concepts and theoretical unification'. *Behavioural and Brain Sciences* 33(2/3): 219-220
- Martin, M.G.F. 1997. 'Sense, Reference, and Selective Attention'. *Aristotelian Society Supplementary Volume* 71(1): 75-98
- McDowell, J. 1990. 'Peacocke and Evans on Demonstrative Content'. *Mind* 99(394): 255-266
- 1994. *Mind and World*. Cambridge, MA. Harvard University Press
- 1998a. *Meaning, Knowledge, and Reality*. Cambridge, MA.: Harvard University Press
- 1998b. 'In Defence of Modesty', in McDowell 1998a: 87-107
- 1998c. 'Another Plea for Modesty', in McDowell 1998a: 108-131

- 1998d. ‘On the Sense and Reference of a Proper Name’, in McDowell 1998a: 171-198
- Medin, D.L. & Schaffer, M. 1978. ‘A Context Theory of Classification Learning.’ *Psychological Review* 85: 207-238
- Millikan, R. 1984. *Language, Thought and Other Biological Categories*. Cambridge, MA: MIT Press
- 1990. ‘The Myth of the Essential Indexical’. *Noûs* 24(5): 723-724
- 2000. *On Clear and Confused Ideas: An Essay on Substance Concepts*. Cambridge: CUP
- 2011. ‘Loosing the Word–Concept Tie.’ *Aristotelian Society Supplementary Volume* 85(1): 125-143
- 2012. ‘Are There Mental Indexicals and Demonstratives?’ *Philosophical Perspectives* 26: 217-234
- 2013. ‘Afterword’, in D. Ryder, J. Kingsbury & K. Williford (eds.). *Millikan and Her Critics*. Oxford: J. Wiley & Sons
- Mount, A. 2008. ‘Intentions, Gestures, and Salience in Ordinary and Deferred Demonstrative Reference’. *Mind & Language* 23(2): 145–164
- Nickerson, R. 2004. *Cognition and Chance: The Psychology of Probabilistic Reasoning*. Mahwah, NJ: Lawrence Erlbaum Associates
- Ninan, D. 2015. ‘On Recanati’s *Mental Files*’. *Inquiry* 58(4): 368-377
- O’Brien, L. 2011. ‘Ordinary Self-Consciousness’, in J. Liu and J. Perry (eds.), *Consciousness and the Self: New Essays*. Cambridge: CUP: 101-122
- Onofri, A. 2015. ‘Mental Files and Rational Inferences’. *Inquiry* 58(4): 378-392
- Oyserman, D., Elmore, K., & Smith, G. 2011. ‘Self, Self-Concept, and Identity’, in M. Learey and J. Tangney (eds.), *Handbook of Self and Self-Identity*. New York, NY: Guilford: 69-105
- Pagin, P. 2013. ‘The Cognitive Significance of Mental Files’. *Disputatio* 5(36): 133-146
- Papineau, D. 1987. *Reality and Representation*. Oxford: Basil Blackwell
- 2013. ‘Comments on Recanati’s *Mental Files*: Doubts About Indexicality’. *Disputatio* 5(36): 159-176
- Papineau, D. & Shea, N. 2002. ‘Ruth Millikan’s *On Clear and Confused Ideas*’. *Philosophy and Phenomenological Research* 65(2): 453-466
- Peacocke, C. 1981. ‘Demonstrative Thought and Psychological Explanation’. *Synthese* 49(2): 187-217
- 1992. *A Study of Concepts*. Cambridge, MA.: MIT Press

- 1997. 'Concepts without Words', in *Language, Thought, and Logic*, R. Heck (ed.). Oxford: Clarendon Press
- 1998. 'Implicit Conceptions, Understanding, and Rationality'. *Philosophical Issues* 9: 43-88.
- 2000. 'Fodor on Concepts: Philosophical Aspects'. *Mind & Language* 15(2-3), 327-340
- 2003. *The Realm of Reason*. Oxford: OUP
- 2004. 'Interrelations: Concepts, Knowledge, Reference, and Structure'. *Mind & Language* 19(1): 85-98
- 2005. 'Rationale and Maxims in the Theory of Concepts'. *Noûs* 39(1):167-178
- 2008. *Truly Understood*. Oxford: OUP
- Perry, J. 1977. 'Frege on Demonstratives'. *The Philosophical Review* 86(4): 474-97
- 1979. 'The Problem of the Essential Indexical'. *Noûs* 13(4): 3-21
- 1980. 'A Problem About Continued Belief'. *Pacific Philosophical Quarterly* 61(4): 317-332
- 2001. *Reference and Reflexivity*. Stanford: CSLI Publications
- Piccinini, G. & Scott, S. 2006. 'Splitting Concepts'. *Philosophy of Science* 71(2): 390-409
- Pinillos, N.A. 2011. 'Coreference and Meaning'. *Philosophical Studies* 154(2): 301-24
- Pippin, R. 'Leaving Nature Behind,' in N. Smith (ed.), *Reading McDowell*. Routledge: London: 58-75
- Prinz, J. 2002. *Furnishing the Mind*. Cambridge, MA.: MIT Press
- Prinz, J. & Clark, A. 2004. 'Putting Concepts to Work: Some Thoughts for the Twenty-First Century'. *Mind & Language* 14(1): 57-9
- Pryor, J. forthcoming. 'Mental Graphs'. *Review of Philosophy and Psychology*
- Quine, W.V.O. 1960. *Word and Object*. Cambridge, MA.: MIT Press
- Rattan, G. 2013. 'On the Sense and Reference of the Concept of Truth'. *Philosophy* 88(3): 433-450
- Recanati 1993. *Direct Reference: From Language to Thought*. Oxford: Blackwell
- 2012. *Mental Files*. Oxford: OUP
- 2013. 'Replies'. *Disputatio* 5(36): 205-240
- Rey, G. 1985. 'Concepts and Conceptions'. *Cognition* 19(3): 297-03
- 2010. 'Concepts vs. conceptions (again)'. *Behavioural and Brain Sciences* 33(2/3): 221-222
- Rives, B. 2009. 'Concept Cartesianism, Concept Pragmatism, and Frege Cases'. *Philosophical Studies* 144 (2): 211-238

- Rosch, E. & Mervis, C. 1975. 'Family Resemblance: Studies in the Internal Structure of Categories'. *Cognitive Psychology* 7: 573-605
- Russell, B. 1911. 'Knowledge by Acquaintance and Knowledge by Description'. *Proceedings of the Aristotelian Society* 11: 108-128
- Sainsbury, R.M. 2002. *Departing from Frege*. London: Routledge
- Sainsbury, R.M. & Tye, M. *Seven Puzzles of Thought and How to Solve Them*. Oxford: OUP
- Salis, F. 2013. 'Introduction'. *Disputatio* 5(36): i-vi
- Schroeter, L. 2004. 'The Rationalist Foundations of Chalmers's 2-D Semantics'. *Philosophical Studies* 118(1-2): 227-255
- 2012. 'Bootstrapping Our Way to Samesaying'. *Synthese* 189(1): 177-197
- Smithies, D. 2011. 'What is the Role of Consciousness in Demonstrative Thought?' *Journal of Philosophy* 108(1): 5-34
- Sosa, D. 2010. 'The Fine Line'. *Analysis Reviews* 70(2): 347-358
- Speaks, J. 2014. 'Individuating Fregean Sense'. *Canadian Journal of Philosophy* 43(5-6): 634-654
- Stalnaker, R. 2008. *Our Knowledge of the Internal World*. Oxford: OUP
- Strawson, P.F. 1957. 'Propositions, Concepts, and Logical Truths'. *Philosophical Quarterly* 7(26): 15-25
- 1959. *Individuals*. London: Methuen
- 1974. *Subject and Predicate in Logic and Grammar*. London: Ashgate
- Sutton, J. 2004. 'Are Concepts Mental Representations or Abstracta?' *Philosophy and Phenomenological Research* 68(1): 89-108
- Tiffany, E. 2000. 'What is Essential about Indexicals?' *Philosophical Studies* 100(1): 35-50
- Weiskopf, D. 2009. 'Atomism, Pluralism, and Conceptual Content'. *Philosophy and Phenomenological Research*, 79(1): 131-163
- 2010. 'The theoretical indispensability of concepts'. *Behavioural and Brain Sciences* 33(2/3): 228-229
- Williamson, T. 1996. 'Cognitive Homelessness'. *Journal of Philosophy* 93(11): 554-573
- 2000. *Knowledge and its Limits*. Oxford: OUP
- Wittgenstein, L. 1953. *Philosophical Investigations*. G.E.M. Anscombe & R. Rhees (eds.), G.E.M. Anscombe (trans.). Oxford: Blackwell