

Indistinguishable Senses¹

Aidan Gray

University of Illinois at Chicago

Fregeanism and Relationism are competing families of solutions to Frege’s Puzzle, and by extension, competing theories of propositional representation. My aim is to clarify what is at stake between them by characterizing and evaluating a Relationist argument. Relationists claim that it is cognitively possible for distinct token propositional attitudes to be, in a sense, qualitatively indistinguishable - to differ in no intrinsic representational features. The idea of an ‘intrinsic representational feature’ is not, however, made especially clear in the argument. I clarify it here and, having done so, offer reason to doubt the argument. This will put us in a position to draw some lessons about the relation between *object-directed* and *representation-internal* aspects of cognitive significance.

1 FREGEANISM AND RELATIONISM

Call the role that a representation can play in rationalizing explanation its *cognitive significance*; and call the way that a representation portrays properties and relations as distributed over objects its *referential content*. Frege noted that the cognitive significance of a representation is not determined by its referential content. Examples are easy to find. Famously, (1 a) and (1 b) differ in cognitive significance.

- (1) (a) Hesperus is Hesperus
- (b) Hesperus is Phosphorous

Coming to believe (1 b) might be the result of empirical investigation and might prompt a revision of other astronomical beliefs; no empirical evidence is relevant to (1 a), and coming to believe it would not put one in a position to infer anything that one didn’t already believe. But (1 a) and (1 b) have the same referential content: they both represent Venus as standing in the relation of identity to Venus. Frege challenged us to characterize what, beyond the referential content of a representation, determines its cognitive significance.

The Fregean strategy is familiar. Fregeans posit a two-layer theory of representational content. Each representation—atomic or complex—in addition to denoting its referential content, expresses a sense. Senses are “modes of presentation” of objects. Two representations can share referential content while differing in sense. The cognitive significance of a representation is determined by its sense. The cognitive difference between (1 a) and (1 b) is explained by the fact that “Hesperus” and “Phosphorous” differ in sense.

¹[This is the penultimate version of a paper that is forthcoming in *Noûs*] This work was supported by a fellowship from the Institute for the Humanities at the University of Illinois at Chicago. It has benefited from discussion with an audience at the Institut Jean Nicod, and from discussion with Mahrad Almotahari, Rachel Goodman, Richard Heck, Robert May, Michael Murez, Jim Pryor, Francois Recanati, and two anonymous reviewers at *Noûs*.

The trick, for Fregeans, comes in saying what senses *are*—in giving a theory of sense. Though Frege himself did not offer such a theory, in places he suggested that senses were descriptions: that is, that the sense of a representation is a descriptive condition that the agent associates with it and which its referential content uniquely satisfies. Sophisticated descriptivist approach to sense persist (*e.g.* (*Jackson, 1998*) *Chalmers* (2002))). But the structure of Frege’s theory does not require the identification of senses with descriptions. More on this shortly.

Relationism is a less familiar response to Frege’s puzzle.² It rejects the *form* of the Fregean approach, rather than some particular theory of sense. Put simply: where Fregeanism posits representational *properties*; Relationism posits representational *relations*. Relationism aims to explain the cognitive difference between (1 a) and (1 b) in terms of a representational relation which holds between the two occurrences of “Hesperus” in (1 a) but does not hold between the occurrence of “Hesperus” and the occurrence of “Phosphorous” in (1 b).

Relationism generalizes this kind of explanation beyond intra-representation relations to inter-representation relations. For the Relationist, the cognitive significance of a particular representation—a particular sentence or mental state—depends in part on relations which hold between its elements and elements of other representations. The difference in cognitive significance between sentences with only a single occurrence of “Hesperus” or “Phosphorous”—like (2 a) and (2 b)—is explained by the fact that the occurrence of “Hesperus” in (2 a) stands in a representationally-relevant relation to occurrences of “Hesperus” in *other* representations (*mutatis mutandis* for “Phosphorous” in (2 b)).

- (2) (a) Hesperus is visible in the morning
- (b) Phosphorus is visible in the morning

This paper explores the debate between Fregeanism and Relationism by evaluating a style of argument for Relationism. Before continuing, a note: so far I’ve been talking generically about representations. Frege’s puzzle has purchase both with respect to language and with respect to thought. Here I’ll be focusing narrowly on thought, leaving language largely to the side. We will consider Fregeanism and Relationism as competing approaches to the content of propositional attitudes.

2 RELATIONISM

I’ll assume that Fregeanism is familiar. Relationism is less established, so it will be worth introducing it in more detail. The point of contention between Fregeanism and Relationism is the nature and explanatory status of *coordination*. When two singular terms are coordinated, they represent their referential content ‘from the same perspective’.³ In

²As I understand the landscape, the approaches in (*Putnam, 1954*), (*Fodor, 1990*), and (*Fiengo and May, 2006*) might be called *proto-relationist*. I would classify (*Taschek, 1995, 1998*), (*Fine, 2009*), (*Pinillos, 2011, 2015*), (*Heck, 2012*), (*Schroeter, 2012*), and (*Pryor, 2016*) as Relationists proper. For an overview, see (*Gray, 2017*).

³Though I adopt the term ‘coordination’ from Fine (2009), I mean to be using in a theory-neutral way (sometimes Fine uses it in this way, and sometimes he uses it as to refer to an aspect of his Relationist apparatus). Coordination, in this theory neutral sense, just picks out a rationally relevant relation between representations. Taken in this sense, Fregeanism and Relationism offer different theories of coordination.

an informal sense, the coordination of two singular terms reflects the representational *presupposition* that they share referential content. This is to be contrasted with a representation which explicitly represents two objects *to be* identical, like (1 b) (*cf.* Fine’s distinction between representing *as* the same, and representing *to be* the same (*Fine*, 2009, pg. 40)).

Coordination is relevant to cognitive significance in different ways (*cf.* (*Millikan*, 1997) on “mediate inferences”). The fact that singular terms are coordinated is relevant to their joint inferential potential. Inferential transitions whose truth-preserving character depends on the identity of referential content across particular positions in their premises require that those positions are coordinated (*cf.* Campbell’s (1987) *trading on identity*). For example, from *Hesperus is F* and *Hesperus is G* an agent is in a position to infer that *something is F* and *G*. The same inference is not licensed from *Hesperus is F* and *Phosphorous is G*. An analogous point can be made with respect to action; in many cases, that two representations—say a belief and a desire—about the same object can rationalize an action directed at that object depends on coordination across those representations: the belief that Hesperus is visible, and the desire to see Phosphorous don’t jointly rationalize looking up. Finally, evaluations of the rationality of representational states depends on patterns of coordination: believing that Hesperus isn’t Hesperus has a different rational status than believing that Hesperus isn’t Phosphorous.⁴

Coordination is at the heart of Relationist approaches to cognitive significance. The function of the representational relations posited by Relationists is to establish coordination between representations. Relationism can be characterized by two theses, one positive and one negative. The positive thesis is Cognitive Significance as Coordination (CSC).

CSC Differences in cognitive significance between representations with the same referential content are explained by coordination (and its absence).⁵

The idea is that coordination between representations can do the explanatory work that Fregeans want sense to do. Relationists hold that the difference in cognitive significance between two representations that share referential content is explained by differences in the way that elements of those representations are coordinated with elements of other representations.

CSC should be understood more as a programmatic stance for Relationists, rather than a strict maxim. Though this is not always explicit in their work, Relationist should not insist that *every* cognitive difference between referentially-equivalent representations is explained by coordination. I’m not aware, for example, of a Relationist treatment of the cognitive difference between a third-person thought and a first-person thought with the same referential content.⁶ As we will consider it, then, Relationism is a broad

⁴In this paragraph, and elsewhere in the paper, I write as though attitudes attributed with the same name are coordinated and attitudes attributed with different names are not coordinated. This is at best an imperfect generalization. I rely on it in this paper only as an expository device.

⁵See (*Gray*, 2017) for further discussion of this thesis. See also (*Fodor*, 1990), (*Taschek*, 1995), (*Taylor*, 2003), (*Fine*, 2009, esp. Chp. 2–3), (*Heck*, 2012, esp. section 2–3).

⁶Heck (2012, pg. 161) acknowledges that the difference between a first-person thought and a third-person thought might be explained otherwise than by coordination. A reviewer asks whether there might be a Relationist version of the kind of scepticism about the *de se* we find in (*Cappelen and Dever*, 2013) and (*Magidor*, 2015). For our purposes, I will mark this as a good question and register my inclination to think not (for reasons that are related to the discussion below).

approach, which would need to be precisified to make clear predictions. One of the upshots of this investigation will be a constraint on how it should be developed. For now, we can simply note that Relationists have thought that they could explain the cognitive difference between, *e.g.*, (1 a) and (1 b) by appealing only to coordination.

Relationism's negative thesis concerns the representational structure of coordination; we'll call it Relationism about Coordination (RC).

RC The coordination of referential content is not determined by the sameness of (or resemblance between) intrinsic representational features. Coordination is an irreducibly relational representational feature.⁷

Much of the action of the paper will be directed at developing a characterization of the idea of an *intrinsic representational feature*. For now we can introduce the idea the way that Fine does. Fine contrasts intrinsic representational features with relational representational features. Relational representational features are representational features which "concern" relations *between* representations (2009, pg. 22). For example, the synonymy of two representations is a relational representational feature. Intrinsic representational features are non-relational;⁸ they do not concern relations between representations. So, for example, the *reference* of an expression is an intrinsic representational feature in this sense (*ibid*). But intrinsic representational features are not, in principle, restricted to *reference* and related properties. It is clear that Fine considers senses, of the traditional Fregean style, to be intrinsic representational features (more carefully, he thinks they *would* be intrinsic representational features if there *were* any).

Relationism's negative thesis should be understood against the background of the Fregean approach to coordination. Fregeans reject CSC, and so hold that there is more to cognitive significance than reference and coordination. But they need not deny the rational relevance of coordination either. Sense, for Fregeans, encodes a representation's perspective on an object. So two representations share a perspective if and only if they share a sense. This is to say that, for Fregeans, coordination is *sameness* of sense. This approach to coordination gets its substance from a theory of sense. A theory which provides individuation conditions for senses will determine an account of coordination. Independent of a theory of sense, Fregeanism is only a framework for an account of coordination.

Relationists reject that framework; according to them it misrepresents the structure of coordination. Independently of any particular account of sense, this structure presupposes that the coordination of two representations is a matter of them sharing a representational property. For Relationists, coordination is irreducibly relational.⁹ Representations are not coordinated because they are the same, or similar, along some representational dimension. They are coordinated because a representational relation holds between them. The debate between Fregeans and Relationists is about the structure of the distribution of intrinsic and relational representational features.

⁷See (*Taschek*, 1995, pg. 83), (*Fine*, 2009, pg. 42), (*Pinillos*, 2011), (*Heck*, 2012, esp. pg 144-5), (*Schroeter*, 2012), (*Pryor*, 2016).

⁸It's possible to be misled, in this context, by talk of 'intrinsic' representational features. We do not mean: representational features which are *intrinsic* to the states which bear them.

⁹Care should be taken with this claim. The idea is that coordination is not reducible to any other *representational* features. It might be, for all that Relationists have said, that we could give a reductive account of coordination. It's just that this account wouldn't appeal to any representational properties which explain coordination. Contrast this with the Fregean approach; on that approach, coordination is explained by sense, which is, itself, a representational feature.

This initial characterization of intrinsic representational features is purely negative (*viz.*, intrinsic representational features of a singular term are those that do *not* concern its relation to other representations). This makes it less than perfectly clear what Relationism is committed to denying, and thus makes it difficult to evaluate it. We will work our way towards giving a positive characterization of intrinsic representational features—and thus clarifying the choice between Fregeanism and Relationism—by considering a form of argument that Relationists have deployed against Fregeanism.

3 INDISTINGUISHABLE SENSES

The structural difference between Fregean and Relationist approaches to cognitive significance entails different possibilities for the distribution of intrinsic representational features and coordination-relations. On the Fregean picture, if s_1 and s_2 differ in terms of which other representations they are coordinated with (a relational representational difference between them), they must also differ in their intrinsic representational features (because for the Fregean, coordination is determined by sameness/difference of intrinsic features). Relationists deny this. They hold that two states can differ in their coordination relations without differing in their representational *properties*. But the vagueness of the notion of sense—or, perhaps, the proliferation of theories of sense—makes the Fregean picture a moving target. To cut through the muddle, Relationists pursue a strategy which would refute any possible version of Fregeanism.

First, some terminology. I have been talking about ‘representations’ and ‘representational’ features. By ‘representations’ I simply mean attitude states (state tokens or state types — more on which below). By ‘representational’ properties and relations, I have in mind, paradigmatically, semantic or intentional properties. I’ll keep talking in terms of ‘representational’ properties and relations for a few reasons. First, I’m not sure whether non-descriptive senses characterize semantic/intentional properties of attitude states; it strikes me as a largely terminological matter. Second, there are forms of Relationism which explicitly deny that the relations they posit are semantic features of attitude states (*Heck*, 2012), (*Pryor*, 2016); Almotahari (2013) dubs this position Formal Relationism (for more about this position see (*Gray*, 2017)). The dialectic below is independent of the distinction between Formal and Semantic Relationists, so I use ‘representational’ to encompass both sorts of relation. To get a fix on how I’m using the term, note that I’ll generally be assuming that the representational properties and relations of attitude states are those properties and relations which play a role in rationalizing psychological explanation (I say ‘generally’, because this doesn’t straightforwardly apply to the discussion in section (7)). Readers who are only interested in Semantic Relationism, and who are willing to treat non-descriptive senses as intentional features, can substitute ‘intentional’ for ‘representational’.

More terminology: we will say that coordination relations divide an agent’s representation of the same object into *coordinated bodies*. A coordinated body is a collection of the beliefs, desires, suppositions, *etc* of a particular agent at a particular time that represent an object from the same perspective. Assuming that coordination is transitive¹⁰, coordination relations will establish equivalence classes of representations: each of

¹⁰Fine assumes this about the intra-personal mental case. Pinillos (2011) rejects it. It doesn’t effect the dialectic here, but it simplifies the formulations.

these classes is a coordinated body of representations. Strictly speaking, the elements of coordinated bodies are the object-representations that are constituents of propositional representations. For example, in the belief that would be expressed with “Hesperus is brighter than Phosphorus”, only one of the representations of Venus is coordinated with the representation of Venus in the belief that would be expressed with “Hesperus is visible”. I will often allow myself to talk about coordination between propositional states, when it is clear what is at stake. I will say, for example, that two beliefs are coordinated, when strictly what is coordinated is the representations of a given object which are constituents of those beliefs.

Relationists ask us to imagine a cognitive situation in which an agent has distinct coordinated bodies of attitudes about the same object which are isomorphic in the following sense: there is a one-to-one mapping from one body to the other such that *i*) a representation and its image are constituents of propositional states with the same referential content and force and *ii*) a representation and its image share the same intrinsic representational features. Call this an *indistinguishable sense* (I-S) scenario and call the elements of each body matched by such a function *mirror elements*.

Suppose I-S scenarios are cognitively possible. This would be a refutation of Fregeanism. For any pair of coordinated representations, r_1 and r_2 , in one body there will be a mirror-pair of representations r_1' and r_2' in the other body. Given that r_1 and r_1' share their intrinsic representational features, and r_2 and r_2' share their intrinsic features, the Fregean could not explain why r_1 was not coordinated with r_2' and r_1' was not coordinated with r_2 . So each such quartet would violate the Fregean account of coordination.

The central argument against Fregeanism in (*Fine*, 2009) is a purported I-S scenario. Fine writes:

[L]et us imagine a universe which is completely symmetric around someone’s center of vision. Whatever she sees to her left *is* and *looks* qualitatively identical to something she sees on her right (not that she conceptualizes the two sides as “left” and “right” since that would introduce an asymmetry). *Fine* (2009, pg 36.) [...] Imagine that [she] sees Bruce in “double” and, taking, him to be two people, starts to have simultaneous thoughts with identical content about what each of the supposed two people is like [...] It is clear that she is having two sets of singular thoughts about Bruce. It is also clear that the thoughts in each set are coordinated with one another but not with the thoughts of the other set. The Fregean must, therefore, suppose that Bruce is given through one mode of presentation in one set of thoughts and through another mode of presentation in the other set of thoughts. (*ibid.* pg. 71).¹¹

There are two bodies of attitudes about Bruce, identical in their referential content. The set-up of the example precludes the Fregean from pointing to a difference in senses-construed-as-descriptions. Fine goes on to assert that, given the representational properties of the two bodies, the Fregean can point to no “plausible non-descriptive difference” between the representations in each body (*ibid.*). This amounts to the claim that there

¹¹I combine the description of the case from two different parts of the book because the second description directly concerns thought—so is relevant to us—but it depends on details explained in the first presentation.

is no *other kind* of intrinsic representational differences between mirror elements. And if this is right, it doesn't matter what theory of sense we develop. It won't be able to make sense of this case.

Let's grant that there are no referential differences between the bodies of attitudes about Bruce: if there is only an egocentric spatial difference between the way the two Bruce's appear in the thinker's cognition, and the thinker doesn't possess the capacity to think *about* space in egocentric terms, we can grant that she doesn't represent Bruce as having different spatial properties in the two bodies of belief.¹² But it's not clear why Fine thinks that it is obvious that the Fregean cannot point to any relevant non-descriptive intrinsic difference. The possibility of non-descriptive approaches to sense is not seriously considered in (2009).¹³ And once we engage with those approaches, the status of this example as a genuine I-S scenario is thrown into doubt.

Two notes before proceeding. First, it is natural to wonder whether the dispute between the Fregean and the Relationist is about attitude *types* or attitude *tokens*.¹⁴ In fact, it's about both. Relationists share the background assumption that token attitude states are distinct only if they have different representational features (Fine (2009, pg. 75) calls this principle 'doxastic link'). Though it's true, then, that Relationists posit token attitude states which Fregeans do not countenance, they do so by positing representational features which Fregeans do not countenance. Neither faction countenances brutally distinct token attitude states.¹⁵ It's just that, for the Relationist, token attitude states can differ in their representational relations without differing in their representational properties. So I won't make hay from the token/type distinction in what follows. In our terminology, the assumption that Relationists and Fregeans share is that token representations are not distinct unless they have different representational features.

Second, we should note that I-S scenarios are the *nuclear option* for the Relationist argument against Fregeanism. That is, the possibility of I-S scenarios is supposed to sweep away all forms of Fregeanism. If it turned out that I-S scenarios are not possible—as I will argue it does—it would not show that Fregeanism is correct; it would not show that there is some individuation of intrinsic representational features such that sameness and difference of those features determines coordination. So there is a sense in which the possibility of I-S scenarios is a stronger claim against the Fregean than the Relationist really needs. But the issue is dialectically important because the possibility of I-S scenarios is consistent with the basic Relationist structure and inconsistent with the basic Fregean structure.

¹²Sosa (2010, pg. 351) is sceptical of even this claim.

¹³Fine says a little more about non-descriptive approaches to sense (see pg. 37). But he seems to assume an idiosyncratic picture of what non-descriptive approaches might be. Essentially, he seems to be assuming that non-descriptive approaches must be causal-historical, and must involve something like *awareness* of reference-fixing conditions. These assumptions are not made by the kinds of approach I discuss below.

¹⁴Thanks to a reviewer at Noûs for prompting me to clarify this.

¹⁵This is slightly delicate with respect to Formal Relationists. But the point remains: for Formal Relationists, token attitude states are only distinct if they differ in non-semantic representational features (that is, in non-semantic features are which relevant to rationalizing explanation).

As noted above, Fregeanism doesn't require treating senses as descriptions. There is a tradition of non-descriptive Fregeanism that developed, to some extent, as a way to isolate the Fregean program from attacks on descriptivism. This approach has a number of importantly different implementations — *e.g.* (*Dummett*, 1981, esp. chapter 5), (*McDowell*, 1977), (*Evans*, 1982, 1985), (*Peacocke*, 1981, 1983), (*Récanati*, 2012), (*Dickie*, 2015). I'm going to ignore the differences here. I will characterize enough generic features of the non-descriptive approach to put it into dialogue with Relationism.

Evans expresses the non-descriptive ethos:

Frege's idea was that to understand an expression, one must not merely think of the reference that it is the reference, but that one must, in so thinking, think of the reference *in a particular way*. The way in which one must think of the reference of an expression in order to understand it is that expression's *sense*. No substantial, or positive theory of the notion of a way of thinking of something is presupposed by this conception of sense. If the intuitive notion needs to be supplemented, we can appeal to the general idea of an account of what makes it the case that a thought is about the object which it is about; two people will then be thinking of an object in the same way if and only if the account of what makes the one person's thought about that object is the same as the account of what makes the other person's thought about that object. (*Evans*, 1985, pg. 294)

The idea is that we can hold on to the claim that the cognitive significance of a representation is tied to the way its reference is determined while abandoning the implausible idea that, in general, reference is determined descriptively.

How should we think about reference-fixing if not descriptively? Speaking generally, non-descriptive Fregeans appeal to the way that representational states are bound up with information-processing systems (*Evans*, 1982, sect. 5.2). On this model, senses are understood as ways of gaining information *from*, and acting *on*, objects. Token propositional attitudes are situated in an agent's cognitive architecture so as to have connections to systems of perceiving and acting; connections which are not fully determined by the referential content of those representations. For example, that a particular belief *B* is about *o* might be explained in terms of the way that *B* is hooked up with systems of perception and action which connect *B* to *o*. This will manifest itself, for example, in an agent's disposition to take information derived from a specific perceptual relation to *o* as directly relevant to the truth of *B* (*Evans*, 1982, pg. 121) (*Evans*, 1985, pg. 303ff) (*Peacocke*, 1983, pg. 109ff). Other beliefs about *o*—even other beliefs with the same referential content—might be differently situated in the cognitive system.

We should pause to emphasize the 'directness' appealed to in the characterization of these informational dispositions. The idea is that these dispositions are not fully explained by the referential content of the agent's representations of the relevant object. While blindfolded, I might believe based on testimony that there is a blue cube in front of me. And there is surely a sense in which at the moment before my captors take my blindfold off, I am disposed to treat information from the blue cube in front of me as relevant to that belief. After all, when the blindfold comes off I'll see it and if it looks expensive, I'll come to believe that it is expensive. But that disposition depends on my

believing of the object that it is a blue cube. These are not the kind of dispositions that are relevant to non-descriptive senses on this model (Evans, 1982, pg. 271). The idea is that there are dispositions which are *not* grounded in representing objects as bearing properties and then identifying them *via* those properties.¹⁶

As the passage from Evans illustrates, the non-descriptive approach to sense generates an account of coordination (see also (Peacocke, 1981, pg. 195)). Two representations which refer to *o* are coordinated if and only if the account of what makes them about *o* is the same. For this to be of much help, we would have to a sense of how to individuate such accounts. And this isn't straightforward.

To get a sense of the approach, we'll focus for now on 'easy' cases—cases where the non-descriptive account appears to offer illumination. Suppose that an agent has two uncoordinated beliefs about Jane, one which she would express with “Jane Smith is an accomplished mathematician” and the other she would express with “That woman dropped her hat”. It doesn't seem implausible to suggest, as the non-descriptive Fregean does, that the fact that the two representations of Jane are not coordinated is connected to the fact that the two beliefs are differently situated with respect to information-processing. The second belief deploys a way of thinking about Jane which is made possible by a perceptual link to Jane, and could be deployed in an action which requires representing Jane in egocentric space (say, reaching out to return her hat). The first belief deploys a way of thinking about Jane which is made possible by the ability to understand utterances involving the name “Jane Smith” and which could be deployed in name-involving utterances about her.¹⁷

What is relevant for us going forward is that on this family of approaches, the sense of a representation is constitutively connected to ways of perceiving and acting on objects that are not determined by their referential content. Recall that we are trying to understand the way that the cognitive significance of a propositional attitude might exceed its referential content. Non-descriptive Fregeanism holds that propositional attitudes might be such that they stand in rationally-relevant relations to objects which are not determined by referential content. Two beliefs about the same object might differ in the way that they are rationalized by perceptual relations to that object, or might differ in the way that they rationalize actions directed at that object. Here, then, is one thing that an *intrinsic representational feature* might be: a rationally-relevant information-dispositional relation to an object.

5 INDISTINGUISHABLE SENSES AGAIN

With the outlines of non-descriptive Fregeanism in view, let's return to Fine's putative I-S scenario. Note that Fine doesn't explain how it comes to be that the thinker sees Bruce “in double”. But this is a crucial aspect of the case for the non-descriptive Fregean.

¹⁶This means that non-descriptive Fregeans must rely on empirical results about perceptual information-processing in developing their approach. For a representative example, see (Dickie, 2011).

¹⁷A question arises about whether these name-using dispositions can really be fit into the same mold as perceptual relations to objects. It might be, on the contrary, that the relevant informational dispositions are grounded in agents' representations of the properties of the named individuals. I suggest as much in (Gray, 2016). For our purposes, it doesn't matter what kind of direct information-dispositions exist. The issue between the Fregean and Relationist is whether appeal to *any* such dispositions is necessary to characterize cognitive significance.

Suppose that the illusion is the result of well-placed mirror. So the two bodies of belief about Bruce are generated via attention to different parts of the visual field. This is exactly the sort of difference that non-descriptive Fregeans would treat as a difference in sense (see, for example, (Peacocke, 1983, pg. 110)). The two bodies of information are differently related to information derived from the object via different parts of the visual field. These dispositions are maintained by non-conceptual abilities to track objects as they move through egocentric space—see (Scholl, 2001, esp. sect. 5)—and so are ‘direct’ in the relevant sense.

A note of caution. I don’t want to suggest that it is easy to pull off the non-descriptive Fregean story in cases like this. There are real questions about how to individuate information-channels in a way that makes the non-descriptive approach plausible. The point here is just that Fine hasn’t offered us any reason to think it can’t be done. For I-S examples to do their dialectical job, these approaches must be engaged with.

Pryor (2016) re-imagines the I-S scenario, with this kind of worry in mind.¹⁸ He writes:

Suppose Flugh is an alien whose many eyes are on long stalks, which can wriggle through the maze of twisty little passages where he lives. The experiences generated by his eyes are co-conscious, but not always spatially integrated. On one day he has two qualitatively matching experiences of a homogeneous sphere, without any presentation of how the spheres are spatially related to each other. Neither experience even seems to be above, or to the right, of the other. As it turns out Flugh is seeing only a single sphere. In fact, it may be that he’s only seeing a single sphere with a single eye, but through some signaling glitch in his brain, he now has cognitively distinct mental presentations.” (2016, pg. 331-2)

The basic structure of the example is the same. We have distinct isomorphic bodies of representations; they share referential content and corresponding elements appear to differ in no intrinsic representational features. This version has features which make it more troubling for the Fregean. The visual experiences generated by the different eyes don’t combine into a visual field, so it isn’t clear what it comes to to say that they present the sphere from a different perspective. If the perspectives differ, they seem to differ merely numerically. More than that, Pryor floats the idea that the two presentations might derive from the same eye-stalk and their division might be the result of a ‘glitch’; so the Fregean could not even claim that the difference in sense should be located in the relation each stands in to different eye-stalks.

But even here, the non-descriptive Fregean will think that some details have been left out. First, what’s the nature of the ‘glitch’? Is it a one-time thing? If new information comes in through the eye-stalks, where does it go? Suppose that the two bodies of information are such that *going forward* new information derived from different eye-stalks will be taken as differentially relevant to them. If information comes in from stalk 1 it is added to one body; if information comes in from stalk 2, it goes into the other body. If this is the case, the situation is assimilable to the non-descriptive Fregean model. Elements of the two bodies of information differ intrinsically in terms of their dispositional relation to different eye-stalks.

¹⁸Pryor doesn’t explicitly discuss non-descriptive Fregeanism. But it’s clear that he deploys this example to avoid the kind of problems we found with Fine’s example. See his note 35.

So to get a genuine I-S scenario we shouldn't fill out the example that way. Let's suppose that either there is no regularity in how incoming information is sorted, or that if there is a regularity it treats the two bodies of information the same way (perhaps, for example, new information coming in from the relevant eye is 'split', and a representation with the same referential content is added to each body).

Another question: how does Flugh translate his perceptual states into *action*? Suppose that he sees something of interest on one of the visual feeds and wants to take a closer look. What explains why the relevant eye-stalk moves towards the seen object? It's possible, for anything that Pryor has said, that this depends on information-dispositional facts about Flugh's representational states. That is, it might be, as matter of cognitive architecture, that information derived from a particular eye-stalk can govern the movement of that stalk without special mediation by representational states *about* the stalk.

Again, we can ask a question about the 'glitch': do elements of the two bodies which result from the glitch differ in their information-dispositional relations to action? Perhaps one body stands in a direct action-guiding relation to one eye-stalk, and the other body stands in a direct action-guiding relation to another stalk. But, again, this would make the case fit the non-descriptive Fregean mold. So we shouldn't fill out the example that way. We should suppose that the two bodies of information stand in the same primitive action-guiding relation to the movement of each stalk (and to any other action-types Flugh is capable of performing).

With all of this in place, are we sure that I-S scenarios are cognitively possible? Do they characterize a possible arrangement of cognitive properties and relations? Conviction begins to waiver. That purported I-S scenarios involve *perceptual* representations is revealing. Focus on perceptual states can smuggle in the kind of intrinsic representational differences that non-descriptive Fregeans latch onto. Once we clearly stipulate that those sorts of differences are absent—once we stipulate that mirror elements of each body stand in the same information-dispositional relations to any possible perception or action Flugh might involve himself in—it is less clear that I-S scenarios are cognitively possible.

6 INDISTINGUISHABLE SENSES AND RATIONALIZING EXPLANATION

It will be helpful, at this stage, to move to a higher level of abstraction. The kind of response that non-descriptive Fregeans would give to I-S scenarios illustrates a more basic clash between the Fregean and the Relationist.

To get perspective here, we can consider how I-S scenarios interact with rationalizing explanation.¹⁹ If there were an agent with merely relationally distinct bodies of attitudes, what would this mean for the possibilities of explaining her states and behaviours? Let's frame the question this way: consider an agent in an I-S scenario—call her Mira—and an agent who differs from Mira only in the absence of one of the merely relationally distinct bodies of attitudes—call her Una. Are there any rationalizing explanations we can give of Mira which we cannot give of Una? And if so, what is their character?

What would be ideal for the Fregean would be to show that even according to the Relationist's *own lights*, there are no rationalizing explanations which could apply to

¹⁹The approach taken in this section is loosely inspired by the methodology in (*Ninan*, 2016).

Mira and not to Una. But this is not the case. In this section, we will see that merely numerically distinct attitudes give themselves explanatory work to do. But we will also see that the explanatory pay-off is far more meagre than we might have expected. In section (7), I'll suggest the payoff has a structure that makes it difficult to see what we gain by recognizing it. In a sense, one already has to be a committed Relationist to accept that I-S scenarios are possible.

Before comparing Mira and Una, we should focus on Mira and see what having merely relationally distinct bodies of representations means for the possibilities of rationalizing explanation. Before that, some ground-rules:

First, we'll assume that propositional attitudes can figure in the explanans and the explananda of rationalizing explanation. But we won't assume that propositional attitudes are the only sorts of things that figure in explanations. We will allow, for example, that perceptual relations to objects might figure in the explanation of an agent's beliefs; or that an agent's propositional attitudes might explain her actions. This is required to allow space for the non-descriptive Fregean position, which holds that senses are rationally-relevant relations to objects. We will think of the general form of rationalizing explanation as follows:

$$I_1 \dots I_n \rightarrow O$$

Here $I_1 \dots I_n$ are input states; they are some mix of attitudes and relations to objects. And O is an output state: an attitude or an action. $I_1 \dots I_n$ jointly explain O .

Second, we'll assume that rationalizing explanations subsumes token attitude states in virtue of their representational properties and relations. And so rationalizing explanation is counterfactually robust with respect to representational properties and relations: if we have a genuine rationalizing explanation of some outcome O partially in terms of some token attitude states, it follows that other token states which have the same representational properties and stand in the same representational relations would *ceteris paribus* generate a token state which shares O 's representational properties and representational relations to the inputs. I'll make no assumptions about how to cash out the *ceteris paribus*; nothing will turn on it here.

In a Frege case, an agent has distinct attitudes with the same referential content and the same force (*belief, desire, etc*) but which participate in different rationalizing explanations. To take an example friendly to the non-descriptive Fregean: suppose I have a belief that I would express with "Jane Smith wants to shake my hand", call it ' B_{js} ', and a perceptual belief to the effect that the woman in front of me wants to shake my hand, call it ' B_p '. Unbeknownst to me, Jane Smith is the woman in front of me.

B_{js} and B_p can participate in different rationalizing explanations. For example, B_p is part of the explanation of my reaching out my hand towards Jane in my current circumstances (call this action ' R '). Schematically, then, there is a rationalizing explanation of R that appeals, amongst other inputs to B_p :

$$(3) \quad I_1 \dots, B_p, \dots I_n \rightarrow R$$

I make no assumptions about $I_1 \dots I_n$; presumably they involve other beliefs and desires. We can abstract away from the details. The relevant point is that the explanation is such that B_{js} could not occupy the same explanatory role:

$$(4) \quad I_1, \dots, B_{j_s}, \dots, I_n \not\rightarrow R$$

Two notes. First, I don't mean that B_{j_s} could never play any part in explaining R . Just that it couldn't play the part that B_p plays. Recall the non-descriptive Fregean idea that attitudes can stand in rationally relevant relations to ways of acting on objects which are not determined by their referential content. Second, I am not actually endorsing this style of explanation. The point is just to set up a contrast between how Fregeans and Relationists must conceptualize I-S scenarios.

With all of this in place, we can characterize what an I-S scenario would mean for rationalizing explanation. Mirror elements in distinct bodies create Frege cases; they differ with respect to the explanations they can participate in. But given the set-up, they can only differ in a limited way: for any rationalizing explanation which appeals to some set of attitudes, $A_1 \dots A_n$, in one mirror-body there will be a corresponding explanation involving the mirror elements $A_1' \dots A_n'$ in the other body. So while it is true that mirror elements A_1 and A_1' differ in the role they can play in rationalizing explanation, it is also true that wherever substituting A_1' for A_1 changes an explanation into a non-explanation we can change it back into an explanation by further substitutions of mirror-representations.

This claim requires some unpacking. In particular, what counts as a "corresponding explanation" will look slightly different in different cases. We can start with the simplest kind of case. Suppose we explain the fact that Mira forms a belief B_q (with the content that *something is F and G*) by appealing to her beliefs B_1 (with the content that *a is F*) and B_2 (with the content that *a is G*).

$$(5) \quad I_1, \dots, B_1, B_2, \dots, I_n \rightarrow B_q$$

While it is true that substituting *one* of B_1 and B_2 for its mirror elements will yield a non-explanation, substituting them both will yield a genuine explanation.

$$(6) \quad I_1, \dots, B_1', B_2', \dots, I_n \rightarrow B_q$$

By hypothesis, mirror elements differ representationally only in terms of which attitudes they are coordinated with. And so the permuted explanation preserves all of the representationally relevant features of the inputs.²⁰ So Mira's coming to believe that *something is F and G* is over-determined. She would have done the same thing if she only possessed *one* of the mirror bodies.

In (5) and (6) we appeal to elements of the mirror bodies to explain something that happens outside both bodies (in that case, a belief with quantificational content). We will get a structurally similar case if we appeal to elements of a mirror body to explain an action. If elements of one of the mirror bodies explain why Mira reaches out for the cup in front of her, corresponding elements from the other body will explain the same thing. Recall that mirror elements do not differ in their rational relation to action-types. Again we have over-determination.²¹

²⁰This assumes that no other elements of either body are among $I_1 \dots I_n$. If other such elements are among the inputs, we would need to permute them as well. We are also assuming that no coordination relating representations not about a is relevant. For more on this assumption, see the objection in (8.1).

²¹This is slightly delicate. If we individuate actions in part by the propositional states which rationalize them, actions which are rationalized by corresponding mirror elements will be distinct actions. But

The situation is more complicated when we look at explanations in which elements of the mirror bodies appear as the explanandum. Consider a case similar to that of (5) above, but in which Mira forms a belief B_s (with the content that *a is F and G*) in which the representation of *a* is coordinated with the representations of *a* in B_1 and B_2 .

$$(7) \quad I_1, \dots, B_1, B_2, \dots, I_n \rightarrow B_s$$

In this case, merely swapping B_1 and B_2 for B_1' and B_2' will not result in an explanation. The coordination between B_1 , B_2 and B_s is crucial to the explanatory force of (7). Here we must consider an alternate explanation in which the explanandum, B_s , is also swapped for a mirror element, B_s' .

$$(8) \quad I_1, \dots, B_1', B_2', \dots, I_n \rightarrow B_s'$$

If (7) is a good explanation then so is (8). And so if Mira formed B_s then, at least as far as rationalizing explanation goes, she ought to have formed B_s' too. So here we don't have over-determination of single attitude, but the necessity of parallel development of mirror attitudes.

It might be tempting to balk at this point, so I'll pause a moment. It's natural to focus on the role that B_1 and B_2 play in *justifying* the inference which results in B_s . From that perspective, there is no necessity that Mira come to form B_s' simply because she already has B_1' , B_2' . We do not, and indeed could not, make all of the inferences we are in a position to make.

But that is irrelevant to the question at hand. (7) is an explanation of why Mira formed B_s when she did. It contains not only the licensing beliefs but other material as well; presumably $I_1 \dots I_n$ include things like a question or interest which led Mira to draw the relevant inference at the relevant time. For each such item there are two options. Either it does not contain a representation of *a* which is coordinated with the representation of *a* in B_s , in which case, by hypothesis, it will stand in the same explanatory relation to B_s' . Or it does contain a representation of *a* which is coordinated with B_s , in which case, by hypothesis, it will also have a mirror element coordinated with B_s' . The point is simple: whenever we have an explanation appealing to elements of one mirror body, we can swap mirror elements without loss of explanatory power.

We get a similar result if we look at non-inferential additions to each mirror body. Suppose some perceptual episode is part of the explanation of why Mira comes to have a belief in one body of attitudes. The explanation cannot depend on an explanatory relation which some element of that body stands in to the perceptual episode which its mirror fails to stand in. By hypothesis no such relation exists.

So what did we learn about Mira? As far as rationalizing explanation goes, the two mirror-bodies must move in complete lock-step. Elements of one body cannot explain any actions which elements of the other body do not equally explain; they cannot be explained by perceptual relations which mirror elements in the other body are not equally explained by. And if attitudes are added to one of the bodies, the other body will exactly mirror its development.

these actions must be instantiated in the same bodily movements, on pain of breaking the symmetry required by I-S scenarios. So if a mirror-element explains an action, its corresponding element will explain the same bodily movement under a mirror-description. In I-S scenarios, bodily movements explained by appeal to elements of either mirror-body will *also* be explained by appeal to the corresponding elements of the other body: there will be systematic overdetermination.

Back to the contrast between Mira and Una. Recall Una is just like Mira except that she possesses only *one* of the mirror bodies. Any time an attitude is added—via perception or inference—to Una’s single body, two mirror attitudes will be added to Mira’s mirror bodies. Whenever an attitude in Una’s single body explains an action, two mirror attitudes of Mira’s will each explain the same action of Mira’s. This is not a *reductio* of I-S scenarios. But, once we have the situation clearly in view, it isn’t clear that we are describing genuine cognitive differences between Mira and Una. Mira and Una will never differ in what they *do*, and they will never differ in which referential contents they believe; sometimes there will be a referential content that Una believes *once* and Mira believes *twice*. That’s the extent of the difference. It’s hard to see why we would acknowledge this as a genuine cognitive difference unless we were already committed to the Relationist picture of cognitive significance.

We are in a position now to see that the non-descriptive Fregean response to I-S scenarios is an instance of a broader contrast between Fregeanism and Relationism.²² The Fregean insists that if two representations differ in terms of which other representations they are coordinated with they must also differ in intrinsic representational features. Intrinsic representational features are those features which can participate in rationalizing explanation in a way that is not reducible to the rational relevance of coordination. The kinds of object-oriented relations posited by non-descriptive Fregeans—exhibited by (3) and (4)—are only one kind of intrinsic representational feature. Descriptivist forms of Fregeanism would posit structurally-similar explanations. For example, for a descriptivist, a belief B_a (with the content that *Aristotle is wise*) might be part of the explanation of why an agent formed the belief B_t (with the content that *the teacher of Alexander is wise*). According to the descriptive Fregean, another belief with the same referential content might not be able to play the same explanatory role (because the *sense* of the representation of Aristotle in B_a might stand in a constitutive relation to the description *the teacher of Aristotle*). And this failure couldn’t be reversed by further substitutions of coreferential attitudes. From the descriptivist Fregean perspective, coordination is simply not what explains the rational potential of that particular belief. Similar points will apply to any form of Fregeanism which explains senses in terms of *a priori* inferences determined by competence conditions for concepts (for example, (Peacocke, 1992)).

7 INTRINSIC FEATURES IN REPRESENTATIONAL SYSTEMS

I-S scenarios generate a stand-off between the Relationist and the Fregean. Positing merely relationally distinct bodies of representation can only make a cognitive difference to what goes on *within* those bodies, and the bodies will never develop, *via* the mechanisms which rationalizing explanation recognizes, into non-mirrored bodies. Should we count this as a genuine cognitive possibility? It’s hard to know. We can work with a toy example to get some perspective on why we might want to insist that two bodies of representation are not distinct unless some element in them differs in intrinsic representational features.

Imagine that I am part of a joint endeavour which involves various people gathering information, manipulating it, and acting on it for some purpose. Suppose that as part

²²Thanks to Robert May for helpful discussion of this point.

of my place in that project, people give me information and I record it in a notebook. One day, unbeknownst to me, a trickster places a piece of carbon paper in my notebook. When a colleague gives me some information, I write “Venus is visible” in my notebook. The carbon paper does its work, and a duplicate inscription of that sentence is created on another page. Now ask: has the trick with the carbon paper affected the representational state of my notebook? Should we treat the duplicated notebook and a notebook with single inscription, in a counterfactual situation without the carbon paper, as in distinct representational states? Should we think of the duplicated notebook as representing that Venus is visible *twice*, in an uncoordinated way?

I take it the answer is: *it depends*. We haven’t been told enough about the representational system to decide. It doesn’t follow from the fact that there are two spatially distinct inscriptions of the same orthography that the system is in a representational state which differs from the state it would be in if there were only one inscription of that orthography. We could just as easily hold that one representational state is redundantly embodied in two orthographic inscriptions. Imagine instead of the carbon paper prank, I wrote the sentence down in pencil, and then realizing that it would be impermanent, wrote the same message down in pen. Suppose that the two inscriptions overlap in places, but not completely. Or, suppose they don’t overlap at all, but are immediately adjacent? In none of these cases does it automatically follow that the duplicate orthography is representationally relevant.

So let’s imagine different ways of filling in the scenario.

Version 1 is as follows: people give me information, I write it down on a random page of the notebook. Every now and again someone asks me a yes/no question, for example: “Is Venus visible?”. I start at a random page in my notebook, and look at each page in some order. If I find the information that Venus is visible in the notebook I answer ‘yes’ and stop looking. If I don’t find the information after looking at every page, I answer ‘no’. That’s all I do with the notebook, and nobody else has access to it.

If this is the situation, should we treat the carbon-paper prank as altering the representational state of the book? I’m strongly inclined to think not. There is no explanation that one could give of what the system does which would depend on distinguishing the content contributed by the two sentences. At most, the duplication will effect how *quickly* I’m able to perform my job; and it will do this in an unsystematic way (given that I start my search at a random page). The same sort of effect might result from the pencil-and-pen duplication, which might make the inscription difficult to read. The same effect might result from my writing a single inscription in a very spread-out way, or by smudging an inscription to make it difficult to read, or from my being tired on a particular day. I take it that these wouldn’t amount to representational differences.

This claim might seem dogmatic.²³ Someone may wish to insist that there is a notion of representation according to which the duplicated and the unduplicated notebooks are in different representational states. This representational difference would explain, for example, the redundancy of the book’s representation that Venus is visible — note that the duplicated book would continue to represent that Venus is visible through certain kinds of damage to the notebook that would make the unduplicated notebook cease to represent that Venus is visible. There surely is some sense in which this is a representational difference.

²³Thanks to Michael Murez for raising this issue and for helpful discussion of it.

Though I'm sceptical that redundancy of this kind ought to be given an explanation in representational terms, I needn't put my foot down. Two points here. First, positing representational features of this kind is no comfort to the Relationist. Whatever representational notions we must appeal to explain persistence through damage, coordination with other representations will not be among them. The Fregean and the Relationist *agree* about the sorts of explanation in which coordination is relevant (*viz.* trading on identity, action explanation, logical consistence, *etc.*), and explanations of redundancy are not among them. More generally, Fregeanism and Relationism both address themselves to the way that representational features get involved in rationalizing explanation. So the kind of representational explanation on offer here is orthogonal to the debate between Fregeanism and Relationism.

More fundamentally, though, if we acknowledge that representational features can explain redundancy there is no reason not to appeal to *intrinsic* representational features here, not merely to relational ones. If we are convinced that the fact that the duplicated notebook would continue to represent that Venus is visible despite certain kinds of damage means that we should distinguish its representational state from that of the unduplicated notebook, there is no reason not to distinguish the different inscriptions in terms of the particular *kinds* of damage through which they can persist. And this would be a not-merely-numerical representational difference between the two inscriptions. So even on this model of representational features, we wouldn't have posited merely relational representational features.

Returning to our toy scenario, let's consider a different way that the notebook might be embedded in the information-processing system. In version 2, my role in the system is slightly different. Each page of my book is associated with sets of dispositions directed at different colleagues. These dispositions do not rely on information written on the relevant pages (that is, they are 'direct', in the sense from the discussion above). I write down information told to me by the same colleague on the same page, and I only answer a colleague's questions by consulting the relevant page. Each colleague, if they get a 'yes' to their question, goes out and places a wager on the truth of the relevant claim. In this context, the carbon-copy prank will have made it the case that an additional person places a bet on the relevant piece of information. Here I think it's unavoidable to hold that the prank changes the representational content of the notebook. If we are, in general, willing to explain who places what bet in terms of the content of the notebook, it's hard to see why we shouldn't explain the fact that an extra person placed a bet in terms of the fact that the notebook had additional content.

The difference between versions 1 and 2 is located in the fact that in version 2 the carbon-paper prank—a glitch in the system—generates a representation which differs in intrinsic representational features from the original. The duplicate sentence in the second scenario is bound up in a different way with the object-related dispositions of the representational system. In the first case, there is no intrinsic representational difference between the sentence and its duplicate. We might conclude, then, that 'glitches' don't create new representational states unless they alter intrinsic representational features.

The analogy is probably clear at this point, but we haven't yet been completely fair to the Relationist. To get a case that is structurally analogous to I-S scenarios, we need to consider a third version of the scenario. Version 3 is like version 1 in that the different pages are not associated with different dispositional relations to objects. But we need to add that I sometimes draw inferences based on the information in my notebook. To

make the analogy with I-S scenarios complete, we should add three more details. First, that I will only draw inferences based on information represented on the same page, and I always write the conclusions on that page. Second, we must assume that after the prank copies the inscription, the two relevant pages are isomorphic (for every inscription on one page, there is a corresponding inscription on the other). Finally, we must assume that, as a policy, whenever I make an inference on one page I search the rest of the notebook for an analogous inference to make, and I make that inference too. Given that after the duplication, the two pages are isomorphic, my inferential practices will alter each in corresponding ways (recall our discussion of inference in I-S scenarios in the previous section). The two pages will develop in lock-step.

In this case, should we treat the duplication as altering the representational content of the notebook? Are there two merely relationally distinct bodies of representation? It doesn't seem as though the extra complication makes this case relevantly different from version 1. Positing merely relationally distinct representations doesn't allow us to explain anything that we would want to explain by appeal to representational content. Someone could dig in their heels here. But it would be incumbent on them to give us some reason not to treat these cases as only involving the distributed instantiation of a single representational state.

I suspect the analogy with Relationism's description of I-S scenarios is clear. Once we've clarified how the details of I-S scenarios must be filled in (section (5)), and characterized how I-S scenarios would have to interact with rationalizing explanation (section (6)), the claim that I-S scenarios are cognitively possible looks like brute insistence. There is nothing incoherent about the supposition. But clarity about what is being supposed drains the supposition of whatever intuitive support it received from the I-S vignettes. And reflection on a toy example supports the general idea that we should not count bodies of representation as distinct unless they have elements which differ in their intrinsic representational features.

8 OBJECTIONS AND REPLIES

I've argued that we shouldn't acknowledge the possibility of I-S scenarios. I'll respond to objections. Readers who are already convinced can skip to section (9).

8.1 OUTSIDE COORDINATION

You have misrepresented I-S scenarios. You have assumed not only that the mirror bodies are intrinsically isomorphic but also that they do not differ in coordination-relations to any representations outside both bodies. At various points, your argument turned on there being no differences of this kind—see note (20)—so the argument doesn't work.

I have been assuming that. And, in a sense, the objection is correct. But the situation is complicated.

First, we should be clear about what the objector is imagining. Suppose a thinker is in an I-S scenario with respect to an object o : there are two distinct bodies of attitudes about o that are intrinsically isomorphic. But consider a belief in one body with the referential content that o stands in R to a . There must be a belief in the other body with the same referential content, but nothing in the set-up demands that the two

representations of *a* are coordinated with the same other attitudes about *a*. And if those representations of *a* are differently coordinated, this might affect the sorts of explanations which we could give of the subject and thus alter the dialectic outlined above.

To make the possibilities here more concrete, let's return to Flugh. Suppose instead of seeing one sphere in each experience, he sees two (sphere *a* and sphere *b*). Unbeknownst to him he is seeing the same pair of spheres twice. In each body of attitudes he has a belief with the content *a is smaller than b*. Suppose that these two representations of *b* in these distinct beliefs are coordinated with different further representations of *b*. What would this mean for the possibilities for rationalizing explanation about Flugh?

Here we should distinguish two cases. We could have a *double* I-S scenario. That is, the two bodies of information about *b* might be mirrored as well. This is not structurally different than a single I-S scenario and so the arguments developed above will apply.

Suppose, though, that the two bodies of attitude about *b* are not mirrored. Suppose, through some cognitive glitch, one of Flugh's experiences of *b* represents it as blue while the other represents it as green. What would this mean for the possibilities of explaining Flugh's cognitive activities going forward?

First a cautionary point. We must not imagine that this difference is reflected directly in any of Flugh's representations of *a*. He does not, we must insist, represent *a* as *bigger than a blue sphere* in one body of attitudes and as *bigger than a green sphere* in the other body. This would disrupt the symmetry between the bodies of attitudes about *a*. He is *in a position* to draw those conclusions but we must suppose that he has not done so.

Relative to this kind of case, the arguments above will not go through. Elements of the distinct bodies will be able to participate in inferences which don't have corresponding versions in the other body. And the two bodies might develop differently relative to the same inputs. So the kinds of argument we gave in sections (6) and (7) will not apply.

A few points about this objection. First, it is clear that this is *not* the kind of case that Fine and Pryor are imagining. As described, both cases do not involve asymmetries between the mirror bodies of this sort. The case we have been considering is the case that Fine and Pryor have offered us. And this is no accident. There is nothing, from the point of view of the Relationist approach to cognitive significance, that would suggest that these sorts of asymmetries would make any difference to the cognitive possibility of I-S scenarios. It would be mysterious, from the point of view of Relationism, if the possibility of merely relational distinct attitudes depended on intrinsic differences between representations of other objects.

Moreover, I-S scenarios of this kind wouldn't be dialectically effective against all forms of Fregeanism. We can imagine forms of Fregeanism which characterize senses holistically, in terms of a some broad range of cognitive connections in which a representation is embedded. Fregeans of this form would be tempted to find a difference in sense between the two bodies in the kind of cognitive asymmetries imagined here. I don't claim that this is a promising approach. But the point of introducing I-S scenarios was precisely that it was supposed to allow us to avoid getting into the muck with the Fregeans.

8.2 TRANSPARENCY

There is no need for the discussion of rationalizing explanation or for the toy example. The representational properties of attitudes are transparent to thinkers. We can know that I-S scenarios are possible simply by imagining being in one.

I'll grant, for the sake of argument, that representational properties are transparent (though this is tendentious). And I'll grant that if one can first-personally imagine being in a representational state, then it is possible to be in that state. I deny that I can imagine being in an I-S state. I think I can imagine being in the state that Fine describes. But the representational relevance of the difference in perspective in that state is part of what I am imagining, so it's not an I-S state. I don't think I can imagine being in Flugh's cognitive position. The alienness of Flugh's perceptual situation—co-conscious but not spatially-integrated perceptual representations—makes it difficult to know what to imagine.

When trying to think my way into it, I imagine an inner Flugh-homunculus, monitoring a number of different incoming perceptual feeds (imagine the Flugh-homunculus monitoring a range of screens); the relative location of the screens to Flugh-homunculus, gives no information about how the perceptual experiences conveyed by the screens are egocentrically located with respect to Flugh (*i.e.* that the screen is on Flugh-homunculus' left doesn't indicate that the seen object is on Flugh's left; that two screens are adjacent, doesn't indicate that the two seen objects are adjacent, *etc.*).²⁴ But my grip on this scenario—and my grip on the distinctness of the representational states corresponding to the different perceptual channels—is secured by by the imagined *homuncular perspective*. I'm picturing myself as Flugh-homunculus not as Flugh; but then my ability to locate the different screens in egocentric space is crucial to the example, and thus it ceases to be an I-S scenario.

The only grip I have on Flugh's scenario is in terms of third-personal considerations about rationalizing explanation. I don't deny that the first-person perspective is relevant to the individuation of attitude states. I only claim that such a perspective is not available to me, even imaginatively, for a case like Flugh's.

8.3 RAISING THE QUESTION

You've missed a crucial feature of I-S scenarios: in an I-S scenario, the agent can coherently raise a question about whether the objects of each body are identical. This shows that there are two distinct ways of thinking about the object.

Well, I'm not sure why I'm supposed to believe that the agent can raise that question. As far as I can tell, this is simply a version of the previous objection. As before, I can imagine raising the relevant question in Fine's scenario. But my grip on the question depends on the difference in perspective.

²⁴Thanks to Jim Pryor for discussion here.

8.4 META-COGNITION

*You have been focused on what, if anything, the presence of mirror states would explain with respect to a subject's representations of, and interactions with, external objects. But we could ask about something different. We could ask about what a subject would be in a position to think *about their own attitudes* in an I-S scenario. In an I-S scenario, a thinker would be able to consider her two representations of the object and wonder whether they co-refer.*

It is not dialectically effective for Relationists to appeal to meta-cognitive questions in this way in the context of I-S scenarios. If I-S scenarios are dialectically effective against the Fregean, it should be because they provide us with some grip on the first-order difference between having one body of attitudes about an object and having two intrinsically isomorphic bodies of attitudes about an object. The meta-cognitive difference imagined here presupposes that first-order I-S scenarios are real. The objection presupposes that the subject has two distinct first-order attitude states and then proceeds to imagine that she raises a question about them. The existence of such first-order states is exactly what reflection on I-S scenarios was supposed to make plausible.

8.5 NON-IDENTITY

*You've missed a crucial feature of the I-S scenarios. The agent represents the single object *as two*. That is, it is a representational feature of an I-S scenario that the agent believes that the two bodies of information are about different objects. This means that an I-S scenario contains inferential possibilities that the discussion above ignores (for example, for any property F present in both bodies of information, the agent is in a position to infer that there are two F s). Wouldn't this be enough to establish a cognitive difference between Mira and Una?*

First, this is actually a difference between Fine and Pryor's vignettes. Fine says that the agent takes Bruce "to be two people", but Pryor doesn't say one way or another. Pryor's choice is more principled. In general, it is possible to have two ways of thinking about the same object and have no opinion one way or another about the identity of the object thought of in those two ways. Frege puzzles do not depend on representations of non-identity. Two ways of thinking can have different cognitive significance even when the agent has no opinion about the relevant identity. Coordination is the presumption of identity. Lack of coordination is the absence of that presumption, not the presumption of non-identity. Relationists want us to consider merely relational differences in cognitive significance. They thus predict that I-S scenarios should be coherent even in the absence of opinions about identity. Thus I-S scenarios without representations of non-identity are the appropriate test case.

*That doesn't really answer my question. Consider an I-S scenario *with* a representation of non-identity. It is a genuine I-S scenario: the two bodies of representation differ merely relationally. And it avoids the objections associated with the simpler I-S scenario.*

The two issues—the possibility of I-S scenarios with and without representations of non-

identity—are not separable in the way this response suggests. I don't acknowledge the possibility of I-S scenarios, so I don't know how to interpret the instruction to consider an I-S scenario *with* a representation of non-identity.

The position that holds that bare I-S scenarios are impossible, but I-S scenarios which include representations of non-identity are possible is unstable. Suppose I have an I-S scenario with a representation of non-identity. So there two ways of thinking about an object and the objects thought of in those two ways are represented as non-identical. Can I not change my mind about their non-identity, removing the representation of non-identity (but not adding a representation of identity)? If I do that, I'm back to a bare I-S scenario, which we've acknowledged is not possible. Would the idea be that if I remove the representation of non-identity, I'm left with a *single* way of thinking about the object (because bare I-S scenarios collapse into single bodies of representation)? This isn't recognizable as two ways of thinking about an object. It seems more like thinking about a (single) plurality of two objects. We must either hold that both bare I-S scenarios and I-S scenarios with representations of non-identity are possible, or that neither are.

8.6 LANGUAGE

You've ignored the role that language could play in an I-S scenario. In one of Fine's vignettes, the thinker names both seen objects 'Bruce'. The cognitive role of the two bodies can be distinguished in relation to their connection to utterances involving 'Bruce'.

As with the previous worry, the first thing to note is that there is no reason why the Relationist should predict that I-S scenarios require linguistic competence with names for the relevant object. Setting that aside, adding language to the example doesn't introduce features which weren't already present with perception. By hypothesis, the two mirror-bodies stand in the same rationalizing relation any possible outside inputs. This goes for language as much as for perception. If they did not—if we could, say, distinguish two different token names “Bruce”₁ and “Bruce”₂ and hold that each body was related differently to utterances involving the two names—then the two bodies would differ in their intrinsic representational features. So we must hold that for any utterance involving “Bruce” the relevant information must be added to both mirror-bodies or to neither.

Suppose I grant that. It's still the case that an utterance of “Bruce≠Bruce” could be informative to the thinker. This shows that there are distinct bodies of representations.

This response attempts to establish the possibility of bare I-S scenarios in terms of their potential to be transformed into I-S scenarios with representations of non-identity via assertions of non-identity. It presupposes, then, I-S scenarios with representations of non-identity are possible. See the discussion of non-identity above. It also presupposes the subject in an I-S scenarios could understand an utterance about the relevant individual. Though space prevents a discussion here, I raise worries about that assumption in (*Gray*, 2016).

8.7 NON-PROPOSITIONAL REPRESENTATIONS

The entire discussion so far has assumed that coordination can only hold between elements of representations with propositional content. We shouldn't assume that. If we allow that coordination can hold between, say, an element of a belief and an element of a perceptual experience (understood so as to not have propositional content) then we could describe the cognitive distinctness of the two bodies in I-S scenarios.

I *have* been assuming that, though nothing really turns on it. The essential point is this: whatever the domain of possible coordination relations is, I-S scenarios involve two bodies of representations which are merely relationally distinct in that domain and do not differ in any rationally relevant way to anything outside that domain. The essentials of the discussion depended only on those assumptions.

9 THE UPSHOT

Our goal was to clarify what is at stake between Fregean and Relationist approaches to cognitive significance by characterizing the idea of an *intrinsic representational feature* as it appears in Relationism's rejection of Fregeanism. We did that by considering Relationists' *indistinguishable-sense* arguments against Fregeanism. We concluded that intrinsic representational features are those features whose relevance to rationalizing explanation is not reducible to the rational relevance of coordination. And on that basis decided that the I-S arguments do not succeed.

The task of fully unpacking what this result means for the debate between Fregeanism and Relationism must be left for another time. But I will draw three preliminary conclusions, arranged from most secure to most tentative.

First, and most obviously, considerations involving I-S scenarios are not material for a good argument against Fregeanism. One has a reason to hold that I-S scenarios are cognitively possible only if one is already committed to Relationism, thus they are dialectically impotent against the Fregean. Relationists have floated other arguments—for a discussion, see (*Gray*, 2017)—so attention should turn to those.

Second, I'm inclined to think that, independently of the dialectic between the Fregean and the Relationist, the above discussion shows that I-S scenarios are not cognitively possible. If this is right, then Relationists should take care to develop their theories so that they do not entail the possibility of I-S scenarios. In introducing them, we mentioned that the dialectical relevance of I-S scenarios does not lie in the fact that their possibility is entailed by Relationism. It lies, instead, in the fact that their possibility is consistent with the basic Relationist approach and inconsistent with basic Fregean approach. What Relationists will have to do to avoid the unwanted entailment that I-S scenarios are possible is a question we will leave for another time, except to say that it might involve a restriction in the scope of Cognitive Significance as Coordination. It might be that Relationists should acknowledge non-referential intrinsic representational features somewhere in their system.

Third, the notion of intrinsic representational features that we have arrived at is not well-suited to playing the role of *sense* in a traditional Fregean theory. Recall the Fregean idea that facts about coordination are determined by facts about the sameness and difference of sense. First note that the impossibility of I-S scenarios doesn't entail

the Fregean account of coordination. From the fact that there cannot be two distinct bodies of attitudes which are intrinsically isomorphic it does not follow that coordination is determined by sameness and difference of intrinsic representational features.

We made no serious attempt above to offer individuation conditions for intrinsic representational features; we relied only on what would follow, from the perspective of rationalizing explanation, if two distinct bodies of attitudes were intrinsically isomorphic. To seriously pursue the Fregean strategy we would need a theory of the intrinsic representational features of individual representations which determined the conditions under which two representations of the same object are coordinated. Nothing in the above suggests that this is possible.

In fact, we can say something stronger: our discussion gives us materials for an argument *against* the Fregean strategy. Let's return to the toy example explored in section (7). In version 3—in which the different pages were not associated with different informational-dispositions, but I made inferences based on the information on each page—the argument that duplication didn't alter the representational state of the notebook crucially relied on the fact the relevant pages were isomorphic after duplication. If we imagine a version of the case in which duplication doesn't create isomorphic pages, the argument fails to go through. Duplication could, in that case, change which inferences I would be in a position to draw and thus could change how I might answer some question. And we *should* treat this as a representational difference. Call this version 4.

Note that in this version, the original and duplicate inscription do *not* differ in intrinsic representational features. The two pages are not differently situated with respect to the informational-dispositional structure of the system. So version 4 is a counter-example to the Fregean claim that we can't have differences in coordination without differences in intrinsic representational features. The original inscription and its duplicate are coordinated with different representations—in virtue of being on different pages—but they differ in no intrinsic representational features.

Moving back from the toy example to cognition, the literature describes examples of this kind: Paderewski cases.²⁵ In a Paderewski case, a speaker has two distinct bodies of attitudes about the same object and elements of those bodies do not plausibly differ in how they are embedded in the information-dispositional system. Elements in the two bodies will play different roles in rationalizing explanation—they will be deployed in different inferences, they will explain different actions, *etc*—but this will be explained by fact that the distinct bodies contain different referential content (for example, one body contains a belief that Paderewski has musical talent, the other contains a belief that he is a politician).

No doubt this could be questioned. Perhaps Fregeans could characterize senses in some way that allowed us to say that referentially equivalent attitudes in the two Paderewski bodies differed in sense. But I'm sceptical, and the analogy with version 4 of the toy example suggests we shouldn't be optimistic. The impossibility of I-S scenarios establishes only that distinct bodies of attitudes must differ somewhere in their intrinsic representational features. But referential content is an intrinsic representational feature, and so a difference in referential content is enough to secure the distinctness of two bodies.

²⁵For discussion of the relevance of Paderewski cases to the debate between Fregeanism and Relationism, see (*Gray, 2017*).

This is tentative. But it suggests that a plausible approach to cognitive significance will have both Relationist and Fregean elements. It will be Relationist in the sense that it accepts Relationism about Coordination; it accepts that coordination is not determined by other representational features. It will be Fregean to the extent that it accepts that rationalizing explanation is sensitive to informational-dispositional relations to objects that are not determined by referential content.²⁶

REFERENCES

- Almotahari, M. (2013), Flaws of formal relationism, *Thought: A Journal of Philosophy*, 2(4), 367–376.
- Campbell, J. (1987), Is sense transparent?, in *Proceedings of the Aristotelian Society*, vol. 88, pp. 273–292, Aristotelian Society.
- Cappelen, H., and J. Dever (2013), *The inessential indexical: On the philosophical insignificance of perspective and the first person*, OUP Oxford.
- Chalmers, D. J. (2002), Sense and intension, *Philosophical Perspectives*, 16, 135–182.
- Dickie, I. (2011), Visual attention fixes demonstrative reference by eliminating referential luck, *Attention: Philosophical and Psychological Essays*, p. 292.
- Dickie, I. (2015), *Fixing reference*, Oxford University Press.
- Dummett, M. (1981), *Frege: Philosophy of Language*, Harvard University Press.
- Evans, G. (1982), *The Varieties of Reference*, Clarendon Press.
- Evans, G. (1985), Understanding Demonstratives, in *Collected Papers*, pp. 291–321, Clarendon.
- Fiengo, R., and R. May (2006), *De lingua belief*, The MIT Press.
- Fine, K. (2009), *Semantic Relationism*, Wiley-Blackwell.
- Fodor, J. (1990), Substitution arguments and the individuation of beliefs, in *Meaning and method: Essays in honor of Hilary Putnam*, edited by G. Boolos, pp. 63–79, Cambridge University Press.
- Gray, A. (2016), Minimal descriptivism, *Review of Philosophy and Psychology*, 7(2), 343–364.
- Gray, A. (2017), Relational approaches to Frege’s Puzzle, *Philosophy Compass*, .Online First.
- Heck, R. G., Jr. (2012), Solving Frege’s puzzle, *The Journal of Philosophy*, 109, 132–174.
- Jackson, F. (1998), Reference and description revisited, *Noûs*, 32(S12), 201–218.
- Magidor, O. (2015), The myth of the de se, *Philosophical Perspectives*, 29(1), 249–283.
- McDowell, J. (1977), On the sense and reference of a proper name, *Mind*, 86(342), 159–185.
- Millikan, R. G. (1997), Images of identity: In search of modes of presentation, *Mind*, 106(423), 499–519.
- Ninan, D. (2016), What is the problem of de se attitudes?, in *About Oneself: De Se Attitudes and Communication*, edited by S. T. M. Garcia-Carpintero, pp. 86–120, Oxford University Press.
- Peacocke, C. (1981), Demonstrative thought and psychological explanation, *Synthese*, 49(2), 187–217.
- Peacocke, C. (1983), *Sense and content: Experience, thought and their relation*, Oxford: Oxford University Press.
- Peacocke, C. (1992), *A study of concepts.*, The MIT Press.
- Pinillos, N. Á. (2011), Coreference and meaning, *Philosophical Studies*, 154(2), 301–324.
- Pinillos, N. Á. (2015), Millianism, relationism, and attitude ascriptions, in *On Reference*, edited by A. Bianchi, pp. 322–334, Oxford.
- Pryor, J. (2016), Mental graphs, *Review of Philosophy and Psychology*, 7(2), 309–341.
- Putnam, H. (1954), Synonymity, and the analysis of belief sentences, *Analysis*, 14(5), 114–122.
- Récanati, F. (2012), *Mental files*, Oxford University Press.
- Scholl, B. J. (2001), Objects and attention: The state of the art, *Cognition*, 80(1), 1–46.
- Schroeter, L. (2012), Bootstrapping our way to samesaying, *Synthese*, 189(1), 177–197.
- Sosa, D. (2010), The Fine line, *Analysis Reviews*, 70(2), 347–358.
- Taschek, W. W. (1995), Belief, substitution, and logical structure, *Noûs*, 29(1), 71–95.
- Taschek, W. W. (1998), On ascribing beliefs: Content in context, *The Journal of philosophy*, 95(7), 323–353.
- Taylor, K. A. (2003), What’s in a Name?, in *Reference and the Rational Mind*, pp. 1–32, CSLI Publications.

²⁶More speculation: on a plausible interpretation, this mix of features can be found in, for example, in (Récanati, 2012) and (Dickie, 2015).