

# MENTAL FILING<sup>1</sup>

[Forthcoming in Noûs. Please cite the published version]

Rachel Goodman

University of Illinois at Chicago

Aidan Gray

University of Illinois at Chicago

The mental files framework has been a persistent presence in philosophy of mind and language for the last half-century and has enjoyed increased popularity in recent years.<sup>2</sup> It has not gone unnoticed, however, that apparent metaphors -- of files, of information being *contained* in files, etc. -- play a central role in explications of the view. But it has been unclear how seriously we ought to take these metaphors. We aim to characterize the basic explanatory commitments of the approach in non-metaphorical terms. The guiding question for the paper is whether, once we move beyond the metaphors, there is any theoretical role for files. Our suggestion is that there is not. To put our view in a slogan: so called ‘mental file’ theory is committed to mental *filing* but not to mental files.

In particular, we show how to replace the file-metaphor with two theses: one semantic and one metasemantic. We argue that the metaphor of mental files can be cashed out in terms of relational representational facts (*viz.* facts about the coordination of mental representations) and a metasemantic thesis about the role that information-relations to objects play in grounding coordination.

---

<sup>1</sup> The authors contributed equally to this paper. Parts of it were presented at St. Andrews, UIC, and Leeds. The authors would like to thank those audiences. The authors would also like to thank Robbie Williams for comments on an early version of some of the material, Mahrad Almotahari for comments on the paper, and Francois Recanati for many instructive conversations about the issues.

<sup>2</sup> See, *e.g.*, (Grice, 1969), (Donnellan, 1978), (Perry, 1980), (Evans, 1985), (Jeshion, 2002, 2009, 2010), (Recanati, 2012, 2016), (Goodman, 2016a), (Goodman & Genone, 2020) .

More broadly, we also hope to respond to a certain kind of worry about the file framework. In our experience, people are often uncertain about what the mental files framework is supposed to be a theory *of*: is it a philosophical theory of *content* -- as suggested by its being presented as a kind of Fregeanism?<sup>3</sup> Or is it a speculative theory of the functional implementation of the attitudes -- as suggested by appeals to cognitive science and claims about cognitive architecture sometimes used to motivate it?<sup>4</sup> Our suggestion is that it is *both*. But seeing each aspect of the framework clearly requires unpacking the file-metaphor.

## 1 The Mental Files Framework

We are concerned here with what we think of as *the philosopher's notion* of a mental file. That is, we are interested in appeals to mental files that are used to theorise the content of propositional attitudes (we'll say more about the standard commitments of these appeals below). Insofar as files play a role in visual psychology and linguistics, these notions are not our target. We focus here on the account of files recently developed by Recanati (as well as elaborations, modifications and criticisms of that approach) because it is the clearest, most comprehensive picture of mental files as used by philosophers to theorise the content of the attitudes.

The file metaphor has gained substantial traction in philosophy of mind. Given that this paper is about how to interpret the file framework, we cannot put all our cards on the table yet. Our initial presentation of the framework will reflect the ambiguity that pervades most discussions of it.

File theorists deploy the following model in theorizing about mental reference: a subject's thought about objects<sup>5</sup> -- paradigmatically propositional attitudes about objects-- is organized into mental files. Each file contains a collection of mental predicates. That two predicates are contained in the same file reflects the fact that they are taken by the subject to be coinstantiated. Each file (that is in good order) has a referent. The referent of a file is, in the

---

<sup>3</sup> See (Recanati, 2012, chp. 3, 2016, preface). For an expression of this uncertainty, see (Ninan, 2015, p.369 fn 10).

<sup>4</sup> See (Jeshion, 2009, 2010), (Recanati, 2012, p. 29), (Green & Quilty-Dunn, 2017), (Murez, Smortchkova & Strickland, 2020).

<sup>5</sup> Following the tradition, we ignore thoughts about properties.

normal case, determined non-satisfactionally<sup>6</sup>; that is, the file need not refer to an object that satisfies all (or most, or perhaps any) of its contained predicates. Instead, its reference is determined in a broadly causal-historical way. The presence of a predicate in a file corresponds to a doxastic state of the agent (principally, a belief). So if an agent has a file about *o*, containing the predicate '*F*', this reflects, or perhaps just is, her belief that *o is F*. The subject uses her system of files to keep track of objects. If things are going well, she will add new information to a file about *o* only if the information derives from *o* via an *epistemically rewarding relation* to it.

<sup>7</sup> A subject can use information that is stored in the same file in a way that reflects the presumed coreference of co-filed information, for example in inferences that “trade on identity” (more on which below).

Centrally, files are supposed to play a role in theorizing Frege’s Puzzle. A subject is in a Frege case when she has distinct files on the same object. Recanati thinks of the file theory as a kind of non-descriptive Fregeanism: files play the role of senses, but file-individuation is not determined by predicates contained in files. So, following Kripke (on Cicero/Tully), Recanati holds it is possible for distinct files to contain all of the same predicates (2012, p. 40). The claim that files are not individuated by their contained predicates is supposed to hold both synchronically (2012, chp. 3) and diachronically (2012, chp. 7). A file can persist through additions and modifications (perhaps even wholesale) to its contained predicates.

The previous two paragraphs are a standard introduction to the file-framework (especially of (Recanati 2012)). There are subtleties, complications, developments, etc -- some of which we discuss below -- but this is the standard way of characterising the heart of the framework. Central to this framework is the idea of a mental particular--a file--that *contains* predicates, which are sourced and sorted through *epistemically rewarding relations* to its referent.<sup>8</sup> We want to ask: how should we interpret talk of ‘files’ as it occurs here? There are two salient options: we could think of files as theoretical posits, implicitly defined by their role in the theory. Or we could think of files as useful metaphors -- expository devices that are

---

<sup>6</sup> Goodman (2016a; 2016b) explores the place of ‘descriptive’ files in the file-framework.

<sup>7</sup> We’ll say more about epistemically rewarding relations, and their role in the file theory later in Section 2.

<sup>8</sup> Views that do not appeal to the containment metaphor and the idea of information sorted through epistemically rewarding relations are not file theories *in our sense*.

useful in laying out the structure of a theory but reference to which will not remain in its *canonical statement*. The literature tends toward the first route. We claim, though, that the second is preferable. Put roughly, we argue that, in its canonical statement, the ‘file’-theory makes reference to a certain kind of relational representational feature, and a certain kind of mental activity. Mental files need not come into it. In short, we posit mental *filing* without mental *files*.

A final word about aims. Our goal is not to argue *for* the file-framework (about which each of us is ambivalent). We’re rather assuming that the framework has enough going for it that it is worth identifying the most promising version. Our interpretation avoids awkward problems that arise on the standard interpretation and clarifies the explanatory commitments of the theory.

## 2 Individuation and Containment

The standard interpretation of the file-framework treats files as theoretical posits. Recanati holds that files are “mental particulars” (2012, p. 38; 2016, preface ; forthcoming). This reflects, in part, the fact that claims about the *individuation* of files and containment of information in files are taken to be explanatory in the theory. We saw above that in order for the theory to embody a non-descriptive solution to Frege’s Puzzle, Recanati holds that files are not individuated by the information they contain. And given the role that files play in relation to Frege’s Puzzle, the fact that two pieces of information are contained in the same file must have some downstream explanatory import. Exactly *what* that explanatory import is, and whether, once we see it clearly, appeal to files is dispensable, is the topic of the rest of the paper.

That two pieces of information are contained in the same file is supposed to explain the fact that certain rational processes apply to them. Paradigmatically, file theorists talk about a certain kind of inference. There is a class of inferences -- for example, from  $Fa$  and  $Ga$  to  $\exists x(Fx \ \& \ Gx)$  -- whose rational credentials depend on the coreference of the occurrence of terms in the premises. Campbell (1987) noted that the rationality of these inferences requires not only that the relevant terms are in fact coreferential, but that their coreferentiality is encoded in the

premises. Talk of ‘encoding coreference’, here, is supposed to contrast with the presence of an (even implicit) identity premise linking the relevant object-representations.

In the literature, these are known as inferences that ‘trade on identity’. How, exactly, we ought to think about trading on identity will be a theme in what follows.<sup>9</sup> But we will raise an initial complication. It is not made explicit, in the literature, exactly what cognitive kind ‘trading on identity’ is supposed to pick out. When it is discussed, we are almost exclusively given examples that involve a narrow class of deductive inferences -- inferences that, if modelled in a formal language, would involve uniform substitution of an individual constant. But it seems clear that the natural kind here must be broader.

Consider, to give one example, a case of practical reasoning, involving the interaction between belief and desire. I believe that Hesperus is visible and I want to see Hesperus. I form the intention to look towards the heavens. Does that process of intention-formation trade on the identity of Venus in my belief and desire? Clearly it does. As in the inferential case, there is a clear contrast between this case and another kind of case: I believe that Hesperus is visible and I want to see Phosphorus; I only form the intention to look towards the heavens because I believe that Hesperus is Phosphorus. So it looks like ‘trading on identity’ must capture rational transitions other than inference. But we will leave those complications aside for now.

The file-theorist holds that the possibilities for trading on identity depend on information being contained in the same file (Recanati, 2012, p. 49; 2016, p. 23). It will be useful, going forward, to work with a slightly artificial statement of this view:

Mental Files Thesis (MFT): If  $S$  has an attitude with the referential content  $\varphi(a)$  in virtue of the existence of a token predicate ‘ $\varphi$ ’ and an attitude with the referential content  $\psi(a)$  in virtue of the existence of a token predicate ‘ $\psi$ ’,  $S$  can trade on the identity of  $a$  in those attitudes if and only if there is a mental file  $F$  such that ‘ $\varphi$ ’ is contained in  $F$  and ‘ $\psi$ ’ is contained in  $F$ .

Some clarification about the idea of a ‘token predicate’: Recall that files ‘contain’ predicates. If I believe both that  $a$  is red and that  $b$  is red, this corresponds to the presence of a token of a mental predicate for *redness* being contained in a file about  $a$  and a token of a

---

<sup>9</sup> For a critical discussion of this question, see Goodman (manuscript).

predicate for *redness* being contained in a file about *b*. So MFT holds that a subject can trade on the identity of the referent of two beliefs if and only if the token predicates responsible for those beliefs are contained in the same file.

Note, first, that MFT quantifies over files and thus reflects theoretical commitment to them. Second, we have intentionally introduced an imprecision into MFT, which we believe is implicit in most discussions of the file-theory. That is, it is not yet clear what the modal force of ‘can’ is. Does it express a kind of psychological ability? A kind of rational permission? Something else? What, precisely, is the explanatory import of sameness of file in relation to trading on identity? We will offer an account of this below. We start with MFT because we believe the imprecision captures a genuine ambivalence in standard expositions of the file framework.<sup>10</sup>

What MFT does capture is the fact that treating files as theoretically-positing mental particulars confers an explanatory role on file-containment and file-individuation. A certain rational process is to be explained in terms of the fact that two pieces of information are contained in the same file. Below we discuss some hand-wringing about file individuation. But, first, note that talk of *containment* also raises awkward questions.

File theorists say that if I believe that *a* is *G*, this mental state consists in a mental predicate for *G*-ness being contained in a file about *a*. But they also say that files are mental singular terms (or, perhaps, senses). So, my file about *a* is a constituent of my belief that *a* is *G* (or a constituent of its Fregean content). So we have a potentially troubling form of reciprocal containment (Woodfield, 1991, p. 549). Similarly, how should the file theorist understand my belief that *a* stands in *R* to *b*? Is it the presence of the predicate ‘*x* stands in *R* to *b*’ in the *a*-file? Or the predicate ‘*a* stands in *R* to *x*’ in the *b*-file? Or both? If both, are these two independent states?<sup>11</sup> Finally, we saw above that trading on identity can bridge attitudes of different types (e.g. belief and desire). This also complicates talk of containment: we can’t equate containment

---

<sup>10</sup> We didn’t invent this way of talking out of thin air. This is one of the ways that file-theorists describe the connection. Recanati (2016, pg. 23), for example, writes: “If there are two predicates F and G in the subject’s file for a given individual (or in one of his files if he has several), the predicates are coordinated and the subject can infer that there is an *x* which is both F and G.” Prosser (2019, pg. 3) writes: “By definition, then, two tokens are of the same file [...] just if it is possible for the subject to trade on identity between them”.

<sup>11</sup> Recanati (2012, pg. 50) discusses this worry attributing it to Goodsell.

with belief if we want trading on identity to always be explained by containment.<sup>12</sup> For example, if the move from me believing that Hesperus is visible and desiring to see Hesperus to me looking toward the sky involves trading on identity, it's unclear what explanation in terms of co-containment of information in the same file is available to explain the encoding of co-reference in this case. File-theorists must, at least, distinguish different kinds of containment or distinguish 'parts' of a file that correspond to different attitude types.

Abstractly, we take these problems to stem from a single source: the awkward relation between file-structure and belief (Or some other kind of doxastic state). On the one hand, file-containment is supposed to have a quite general explanatory role: files play the role of *sense* or *concept*, and thus should be perfectly general with respect to content (unary vs. relational predication) and attitude type (belief, desire, supposition, *etc*). On the other, the file framework sees *belief* as having a privileged status: insofar as files are clusters of beliefs about their objects, beliefs are the fundamental building blocks of the cognitive structure that explains trading on identity.<sup>13</sup> The containment puzzles illustrate that the file-metaphor cannot easily accommodate both of these ideas at once. Below, we show that eliminating files from the framework allows us to recover both of these ideas without the awkwardness induced by reliance on containment.<sup>14</sup>

File-theorists have responded to some of these worries.<sup>15</sup> We will not survey and evaluate those responses here. Overall, we suspect that many styles of response will engage the file theorist in ultimately unsatisfying epicyclic complications of their theory. But, more importantly for current purposes, we suspect that a satisfying response will in fact converge on

---

<sup>12</sup> See (Ninan, 2015, p. 375-6) for a version of this worry. File theorists could back away from this claim. Our assertion is not that there is no way to resolve these difficulties within the file framework but rather that the difficulties won't arise, and the (sometimes unsatisfying) moves that address them won't be required, given the view we propose later in the paper.

<sup>13</sup> As Recanati (2012, p. 38) points out, Taylor (2003) takes issue with the idea that concept possession supervenes on belief in the file-framework. See (Taylor, 20013, p. 77-9).

<sup>14</sup> Ninan (2015) asks why there is an asymmetry between belief and other attitude-types in the file picture (implying perhaps that the privileged role for belief is unmotivated). We preserve the asymmetry (and seek to answer the question of how it is motivated) in our account of the role of ER relations with respect to coordination relations. We think the problem lies with the containment metaphor.

<sup>15</sup> For example, Recanati (2012, pg. 50) discusses the issue of relational information and the reciprocal containment worry (p. 39-40 n12).

the kind of approach we develop below: one that cashes out the file metaphor leaving no real role for mental files.

Another response to the containment problems would simply be to abandon the file-picture wholesale and adopt one of the many other theories of concepts available. We don't take that route because we think the file-picture contains important insights; what those insights are, and how they can be maintained while dropping the problematic appeal to containment, will emerge as we proceed. We see the containment puzzles not as a sign that the file-picture is on the wrong track, but rather, as a symptom of the fact that file-theorists have not clearly distinguished the *metaphorical* from the *theoretical*. Our goal is to improve the file-theory not completely abandon it.

For now, though, the important point is that the problems discussed above are entirely forced on us by treating files as theoretical posits. If we thought of them as helpful metaphors, but ultimately eliminable from the canonical statement of the theory, these questions simply would not arise.

So far we have seen a *negative claim* about file-individuation (*i.e.* files are not individuated by their content) and a *theoretical role* for file-individuation (*i.e.* sameness of file explains trading on identity). These are close to non-negotiable for file-theorists. What's more controversial is what can be said *positively* about file-individuation.

It is central to Recanati's system that each file is governed by an *epistemically rewarding* (ER) relation to its object. The idea of an ER relation is not made particularly precise (we'll have more to say about this in section (6)). Roughly, it is a relation in which a subject stands to an object that allows the subject to acquire information from it. Perception is a paradigmatic ER relation. Recanati also mentions recognitional capacities, memory, and competence with a proper name (2012).

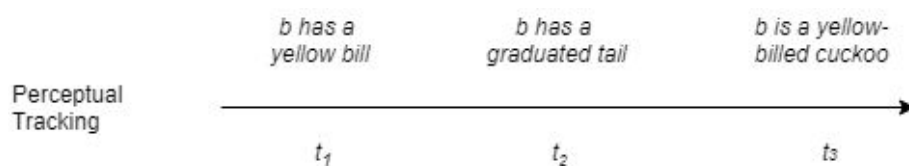
What is important for now is that Recanati (2012) conceived of ER relations as the key to his positive account of file-individuation, both synchronically and diachronically: two files are the same when they are governed by the same ER relation.



Synchronic File-individuation: At  $t$ ,  $F = F'$  if and only if the ER relation that governs  $F =$  the ER relation that governs  $F'$

Diachronic File-individuation:  $F$  at  $t = F'$  at  $t'$  if and only if the ER relation that governs  $F$  at  $t =$  the ER relation that governs  $F'$  at  $t'$ .

We should note, at the outset, that there is a principled, though vague, connection between MFT and this positive proposal (we'll make this connection more determinate in Section 6). There is a connection between the idea that sameness of file is relevant to a certain kind of rational processes, and the idea that sameness of file presupposes sameness of ER relation. Here is the kind of case that one finds in the literature: I'm watching a bird,  $b$ , fly across the sky, tracking it from my left to my right. At  $t_1$ , I judge that  $b$  has a yellow bill. At  $t_2$ , I judge that  $b$  has a graduated tail. At  $t_3$ , because I know that yellow-billed cuckoos have graduated tails and yellow bills, I judge that  $b$  is a yellow-billed cuckoo.

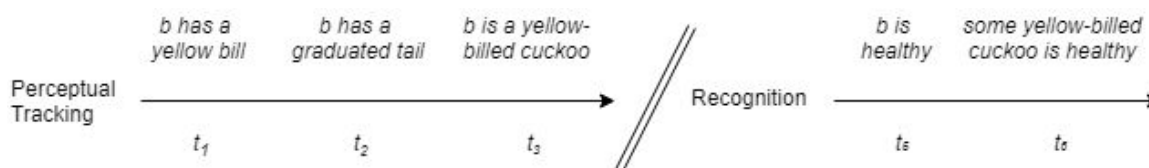


The inference I perform at  $t_3$  trades on identity: the evidence I have that  $b$  is a yellow-billed cuckoo depends on the fact that I have encoded *has a yellow-bill* and *has a graduated tail* as co-instantiated. So we must hold that those two pieces of information were stored in the same file. And the fact that the inference I make at  $t_3$  is in good-standing is surely connected to the fact that I have collected the two pieces of information through a single cognitive relation -- an episode of perceptual tracking -- whose deliverances typically concern a single object. The mental files picture bundles this up in a neat package: trading on identity requires sameness of file; sameness of file requires a continuing ER relation.

We will be offering the file-theorist a way, essentially, to endorse this package. But we argue that once we tidy it up, there's no need to mention files at all.

### 3 Trading on Identity and Diachronicity

We will work our way towards eliminating files by rehearsing a wrinkle in the picture characterized above. After the publication of (Recanati, 2012), a number of critics raised a worry about the diachronic individuation of files--in particular, Recanati's positive proposal that files are individuated by their ER relations.<sup>16</sup> Put simply, the worry goes like this. The simple case above can be easily extended. Suppose the perceptual episode ends at  $t_4$ ; I stop tracking the bird and go home. A week later, at  $t_5$ , I return to the same field and recognize the bird again. I judge that  $b$  is healthy. At  $t_6$  I infer that some yellow-billed cuckoo is healthy.



Plausibly, the inference I perform at  $t_6$  trades on the identity of  $b$ . Even if the recognition of  $b$  at  $t_5$  involves an identity judgment, there is no reason to think that the inference I perform at  $t_6$  is of a fundamentally different character than the one I performed at  $t_3$ . But now there is a problem. The inference at  $t_6$  trades on the identity of information that was acquired through two distinct ER relations: a perceptual relation and a recognitional capacity. So according to Recanati's positive individuation proposal, the two pieces of information are contained in distinct files. And so they should not be eligible for trading on identity.

Moves can be made here. For a start, it's not absolutely clear that we must hold that the relevant ER relation changes from  $t_3$  to  $t_4$ . After all, we have said nothing precise about the individuation of ER relations. But it's hard to avoid the idea that whatever we say about ER relations, some structurally analogous case will be possible (consider a situation in which I learn

<sup>16</sup> See (Ninan, 2015, p. 369-73), (Onofri, 2015, p. 380-4), (Papineau, 2013, p. 167-70). Also, see (Ball, 2015) for the claim that Recanati's files are temporally too fine-grained, but Ball's claim is based on a different kind of argument.

a name for an individual I could previously only perceptually recognize). And so it looks like Recanati's simple picture is in trouble.

Note that Recanati was always aware that the claim that the diachronic continuity of a file was tied to an ER relation needed to be nuanced. He was aware of the kind of everyday case described above. It's obvious that we cannot hold that whenever an ER relation ends, the cluster of information associated with it is simply obliterated. Recanati holds that in such cases a file is 'converted' into a new file, governed by a new ER relation, that contains the information from the pre-conversion file (2012, chp. 7 & 2015, p.74). But this doesn't, by itself, answer the objection -- conversion is not identity.<sup>17</sup>

Recanati (2015, 2016) answers the objection by clarifying that he never meant MFT to be interpreted diachronically. What is important for the mental file framework, he claims, is simply whether the information *is a yellow-billed cuckoo* and *is healthy* is contained in the same file *at*  $t_c$ . This is when the inference takes place, so this is when co-filing matters. It is not relevant that the two pieces of information were *initially* entered into different files. If they find themselves in the same file at some later stage--the stage at which the inference occurs--they stand in the relation relevant to trading on identity.

Here is how Recanati motivates this point. He acknowledges that trading on identity is an inferential process, and thus occurs in time. It is therefore tempting to think that we require diachronic individuation of files to give a file-theoretic account of it. But the question that the mental files framework addresses is the *logical status* of such inferences. And "For logic purposes, a train of reasoning has to be construed as synchronic" (2016, 77). It is therefore appropriate to treat trading on identity as synchronic. Note that when Recanati talks about the "logical status" of the inference, he means, presumably, the fact that the inference is rationally permitted when it is made.

Taking these points on board gives us a revised version of MFT:

MFT-Synchronic (MFT-S): If, at  $t$ ,  $S$  has an attitude with the referential content  $\varphi(a)$  in virtue of the existence of a token predicate ' $\varphi$ ' and an attitude with the referential content  $\psi(a)$  in virtue of the existence of a token predicate ' $\psi$ ',  $S$  is

---

<sup>17</sup> See (Onofri, 2015, 387-88) for essentially this point.

rationaly permitted at  $t$  to trade on the identity of  $a$  in those attitudes if and only if there is a mental file  $F$  such that ' $\varphi$ ' is contained in  $F$  at  $t$  and ' $\psi$ ' is contained in  $F$  at  $t$ .

We will offer a replacement for MFT-S in the next section. At this point, we should note a simple consequence of the move from MFT to MFT-S: the *diachronic individuation* of files plays no explanatory role in relation to trading on identity. Recanati's clarification restricts the *rational relevance* of file-individuation to a synchronic matter.

An additional note: the moral of this section is not happily expressed as Recanati expresses it, by saying that logic treats inference as synchronic. That gives the impression that there is something diachronic--inference--that files are in the business of giving an account of, which must be idealized in some way to fit into our theory. This is misleading. The point is better put by saying that co-filedness of information is supposed to explain how an inference that trades on identity is *licensed*; it explains the *rationality* of the inference. In adopting MFT-S, therefore, there is no need to idealise (into something synchronic) inference qua *temporally extended psychological process*. Rather, all we need is to appreciate that the explanandum of MFT-S is a rational status, not a psychological process. This might seem like a subtle correction but it resolves a central ambivalence of standard presentations of the file theory and will be of central importance in the next section.

The best way to state the moral of this section is therefore that MFT should be replaced by MTF-S because co-filedness of information is meant to explain the rational permissibility of inferences that trade on identity--and it is synchronic co-filing that explains this. If, as we suggest in the rest of the paper, we can *also* do away with the *synchronic* relevance of files, we will have presented a theory of mental filing without mental files.

#### 4 Co-filedness as a Representational Relation

Recall from Section 2 that in common statements of MFT the modal force of 'can' in 'can trade on identity' is left unclear. In Section 3, we saw that MFT in fact makes a claim about what subjects are *rationaly permitted* to do, and that this motivates a synchronic version of the view. In this section, we show that rational permission to trade on identity is constituted by a

*relational representational feature* of the agent's attitude state. With this in view, we will be in a position to directly confront the putative explanatory contribution of files.

First, we'll argue that file theorists should think of the rational permission to trade on identity as constituted by a *representational* feature of a subject's attitude state. A straightforward argument by analogy establishes this. Suppose Smith believes that Twain is an author and Jones doesn't believe that Twain is an author (and doesn't believe of anyone else that they are an author). We might note that because Smith believes that Twain is an author, she can conclude that *someone* wrote a book. Jones is not in a similar position. This is entirely non-mysterious. Again, this is the 'can' of rational permission. And here we take it to be obvious that the relevant permission is constituted by the representational features of Smith's belief: it is because the belief has the content that it does that Smith has the relevant permission.

Our claim is that the file-theorist should say the same thing about rational permission to trade on identity. Suppose Smith and Jones both believe of Twain that he is an author and that he is from Connecticut. Suppose that Smith can trade on the identity of Twain in the two beliefs and that Jones cannot (imagine that Smith would express both beliefs with 'Twain' and that Jones would express one belief with 'Twain' and the other with 'Clemens'). The file-theorist's explanation is that Smith's beliefs are co-filed while Jones's are not. For now we're suspending judgment about that. But surely, whatever we want to say about that, file-theorists will admit that Smith and Jones are in different representational states. The difference between them is not merely at some 'lower-level' of explanation (functional, computational, or whatever). The difference between them is that, in Smith's case, the coreference of the two attitudes about Clemens/Twain is representationally encoded.

To deny this would be to hold that the *kind* of explanation that we give when we say that Smith *can* conclude that someone is an author because she believes that Twain is an author is of a radically different sort than when we say that Smith *can* conclude that some Connecticutian is an author because she believes that Twain is an author and that Twain is from Connecticut. But this is bizarre on its face. It is not as though, when we move from 'single-premise' inferences to 'multi-premise' inferences, we transition from explanation by

appeal to content to some other form of explanation. In both cases, the relevant rational permission is explained by Smith's representational state.

Our second claim is that the representational feature that licenses trading on identity is *relational*. Here we need to introduce some background assumptions. We will suppose that an agent's total attitude state can be decomposed, in part, into individual object-representations. This is to say, for example, that if Smith believes that Clemens admires Twain, we can distinguish the two representations of Twain/Clemens that contribute to the content of that belief. And, in general, given a subject's total representational state, we can distinguish the constituent object-representations that partly constitute that state.<sup>18</sup>

There are different ways to make sense of this assumption. If we assume that attitudes have sentence-like mental vehicles, we could distinguish object-representations in terms of the vehicles that contribute them. Alternatively, we could assume that propositional content itself is structured. If an agent's total attitude state can be decomposed into a collection of relations to Russellian propositions then we can distinguish object-representations in terms of occurrences of objects in the Russellian content of an attitude state.<sup>19</sup>

That an agent's representational state decomposes into individual representations is a substantive assumption. If we both denied the existence of structured mental vehicles and held an unstructured view of content, it isn't clear that it would make sense to distinguish, for example, the 'different' representations of Twain/Clemens in the belief that Clemens admires Twain. But the assumption is commonly made and is certainly part of the file-picture. Officially, we'll remain agnostic about whether content is structured, or attitudes have structured vehicles, or both; but we will assume that at least one of those disjuncts is true.

Given this assumption, we can distinguish, following Fine (2007), intrinsic representational features of attitude states from relational ones. It is, for example, an intrinsic

---

<sup>18</sup> This isn't to say it will always be obvious how to decompose a particular attitude state. Consider the belief that Smith would express with "Twain admires himself". Does it decompose into two representations of Twain or only one (along with the representation of a reflexive property)? Nothing we say here presupposes an answer to questions of this sort.

<sup>19</sup> We take this second way of making sense of the assumption from Fine (2007) and discuss it more fully below.

representational fact about an object representation that it refers to Twain.<sup>20</sup> That two object-representations corefer is a representational relation that holds between them.

Given this set-up, it is clear that what licenses trading on identity is a *relational* representational feature. Smith's license to conclude that some Connecticutian is an author is partly constituted by a representational relation that holds between her two representations of Twain. Let's call the representational relation that licenses trading on identity 'coordination'.<sup>21</sup> Putting everything from this section together, we can endorse the following principle:

Coordination as a Representational Relation (CRR): If  $S$  has an attitude  $\Phi$  involving object representation  $a$  and an attitude  $\Psi$  involving object representation  $b$ ,  $S$  is rationally permitted to trade on the identity of  $a$  and  $b$  if and only if  $a$  and  $b$  are *coordinated*.

None of this should be especially controversial for the file-theorist. They themselves posit the existence of such relations and note their relevance to trading on identity.<sup>22</sup> We're belaboring these points because we'd like to know, precisely, where the theoretical need for files comes in. CRR itself doesn't mention files. So file-theorists must think that CRR cannot stand alone (if it could, then we could certainly explain trading on identity without a theory that quantifies over files). Files must come in to ground or explain CRR itself.

As far as we can tell, there are two possibilities as to how files might ground CRR: They could play a semantic role, or a metasemantic role. From our perspective, part of the difficulty in interpreting the file-framework is that file theorists have not properly distinguished these options. Our aim is to do just this. We will then (in section 5) reject the semantic role and argue (in section 6) that a metasemantic grounding of CRR can be satisfied without appeal to files.

---

<sup>20</sup> We follow Fine's terminology here, though it can be misleading. It is important that we don't interpret the idea of "intrinsic" representational features metasemantically. There are, plausibly, no metasemantically intrinsic representational features -- that is, no representational features that a representational vehicle possess independently from its relations to other things. The distinction between intrinsic and relational representational features should be interpreted as drawing a contrast *within* the representational domain itself, not as contrast about what grounds that domain. We return to this in the next section.

<sup>21</sup> The same relation, or at least a closely related relation, is sometimes called 'direct coordination', 'de jure coreference', 'strict coreference', *etc.*

<sup>22</sup> See (Recanati, 2012, Part IV), (Recanati, 2015, Part I) & (Recanati, 2020)

## 5 Coordination as Irreducibly Relational

We are exploring the idea that files play an explanatory role in relation to coordination. In this section, we consider the possibility that coordination requires a representation-internal explanation; that is, that coordination is a representational relation that is grounded in representational properties. One possible role for files in relation to CRR is that file-containment is -- or constitutes -- a representational property that grounds coordination.

To assess this suggestion, we need some more background. In the previous section, we distinguished between intrinsic and relational representational features. We now need a further distinction: between representational relations that are *grounded* in intrinsic representational features and those that are not. Call the latter irreducibly relational representational features.

That two object-representations co-refer is a relational representational feature. But it is grounded in an intrinsic representational feature: the reference of each representation. If we fix the reference facts about an attitude state, we have also fixed the co-reference facts.

One could assume that all representational relations are grounded in intrinsic representational features. But why should we? In general, there is nothing that tells us that the *relations* in some domain are grounded in the intrinsic features in that domain. Consider, for example, the domain of chess-facts. Is the fact that Smith and Jones are playing a game of chess *against each other* grounded in the intrinsic chess-facts about Smith and the intrinsic chess-facts about Jones? It doesn't seem so. Why should things be different in the representational domain?

Here we should put our cards on the table: we will ultimately offer an interpretation of the file-framework that invokes a *relationist* approach to Frege's Puzzle. Relationism holds that coordination is an irreducibly relational representational feature.<sup>23</sup>

---

<sup>23</sup> The clearest expression of this idea that we are aware of are in (Taschek, 1995), (Fine, 2007), and (Heck, 2012). See (Gray, 2017) for further discussion and references.



It's easiest to understand Relationism as a rejection of the Fregean account of coordination. Recall that coordination is the representational encoding of coreference. Frege's Puzzle teaches us that the fact that two representations co-refer doesn't entail that the representational state of which they are a part encodes that they corefer. The Fregean approach to coordination captures the encoding of coreference, when it is present, in a second layer of intrinsic representational features: sense. Each representation has both a sense and a reference, and coordination is *sameness-of-sense*. So, for the Fregean, coordination is a relational representational feature, but it is determined by intrinsic representational features of the object representations that are coordinated: that is, that they each possess a certain *sense*.

Relationism rejects the Fregean approach: coordination is not determined by the intrinsic representational features of an attitude state. We should be clear about what this means. Recall our chess example, and the claim that the fact that Smith and Jones are playing chess against each other is an irreducibly relational chess fact. This isn't the claim that there is no explanation of what makes it the case that Smith and Jones are playing against each other. It is rather the claim that there is no description of how Smith is playing chess (that doesn't mention Jones) and of how Jones is playing chess (that doesn't mention Smith) that entails that Smith and Jones are playing each other.<sup>24</sup>

Similarly, the Relationist claim is not that there is *no* explanation of why two coreferential attitudes are coordinated. It is the claim that there is no combination of representational facts about Smith's belief that Twain is an author (that doesn't mention her representation of Twain as from Connecticut) and of representational facts about her belief that Twain is from Connecticut (that doesn't mention her representation of Twain as an author) that entails that the two representations of Twain are coordinated.

There is, of course, a question about how to understand coordination as an irreducibly relational representational feature. We won't take a stand on that question here. But to help fix

---

<sup>24</sup> This could be denied, of course. Given the right assumptions about chess-vocabulary and the individuation of games of chess, we might be able to hold that the fact Smith is playing Jones is grounded in the fact that Smith is participating in game *G* and the fact that Jones is participating in game *G*. We have no interest in refuting this position (we use chess only as an example). We consider and reject an analogous proposal about the relation between file-containment and coordination below.

ideas, we'll sketch two possibilities. In (2007) and (2010b), Fine offers two ways of understanding coordination, corresponding to the two assumptions about object-representations we canvassed above: mental vehicles and Russellian content.

On one model, we assume that an attitude state has two 'layers' of content. At one level, there is a sequence of Russellian propositions. This layer captures the referential content of the state. At the other level, we have a relation on occurrences of objects in the first layer. This layer captures the coordination relations. The irreducible relationality of coordination is reflected in the fact that the coordination relation on an attitude state is not fixed by the Russellian content of that state (or any other intrinsic representational feature).<sup>25</sup>

The other model does not invoke Russellian content. In it, the guiding idea is that the semantic facts about some body of representations are fixed by the representational 'requirements' that are true of them. So, for example, it might be a representational requirement that a name *a* refers to an object *o*. In this system, coordination is the requirement of coreference: names *a* and *b* are coordinated when it is a semantic requirement that they corefer. The irreducible relationality of coordination is secured by the fact that, given the logic of semantic requirements, that two names are each required to refer to *o* does not entail that they are required to corefer. To apply this approach to attitudes, we must assume that there are mental vehicles, and that they have enough structure to be the subject of semantic requirements of this kind.

Evaluating the Relationist claim is made difficult by the vagueness (or, at least, theory-dependence) of the idea of a 'representational' feature of an attitude state. But our purposes here don't require that we argue for Relationism; they only require that we argue that the file-theorist should accept it. And this is a somewhat easier task.

The question of whether file-theoretic explanations of coordination are Relationist can be clarified by appeal to the metaphor the theory appeals to. Think of a file-cabinet: each file contains a collection of file-cards, with a predicate written on each card. File-cards in the same file are presumed to contain information about a single individual. But *which* file some

---

<sup>25</sup> For a more substantial discussion see (Gray, 2017).

collection of file-cards is contained in – the top file, the middle file, *etc* – need not, itself, be a representationally significant feature of the filing system.<sup>26</sup> The filing cabinet represents two predicates as coinstantiated because they are contained *in the same file*, not because each is contained in some particular file (say, the third file from the top). What makes a representational difference is the pattern of co-filing; one could switch the location of the file-cards in the bottom drawer with the location of the file cards in the top drawer without altering the representational state of the file cabinet.

The claim that file-theoretic explanations are Relationist explanations--not explanations in terms of intrinsic representational properties--is essentially the claim that, for all that file theorists say, the same is true of mental files. Co-filing is representationally significant. But co-filing is not determined by any other representationally significant feature of files.

And, while file theorists have not been clear about this question, we see it as likely that many of them would be willing to accept this relationist claim.<sup>27</sup> Suppose that Smith has an object representation which refers to Mark Twain, to the effect that he is an author, and an object representation which refers to Mark Twain, to the effect that he is a Connecticutian. And imagine further that these representations are co-filed. We imagine the file-theorist would say (as we would) that this *co-filing fact* is a representational fact: after all, the fact of co-filing (vs. not) implies something about what Smith is rationally permitted to do (she is rationally permitted to make certain inferences, act in certain ways, *etc* if there is co-filing, and isn't if not). But imagine now that I tell you these (co-filed) representations are in *this* particular file rather than *that* particular file. Does this 'haecceitic' fact constitute a representational fact about Smith? It doesn't seem to. And, we see no evidence in the literature that the file-theorist thinks it does. Unless the file-theorist argues that this kind of difference is indeed a representational difference, we think it's fair to construe file-theoretic explanations as relationist.

---

<sup>26</sup> Of course we can imagine a system in which the files themselves have representational significance (perhaps we always put information about the tallest person we know in the top file). The point here is that we needn't imagine that system in that way and, so we will argue, the file-theory doesn't work that way.

<sup>27</sup> Prosser (2019) also suggests that file-approach can be understood as a kind of Relationism. His discussion focuses on issues with the diachronic individuation of files.

An important clarification.<sup>28</sup> Our claim is not that there is no representationally significant aspect to files over-and-above the predicates they contain. In Recanati's system, file *type* is representationally significant. For example, that a file is a *self-file*, or a *perceptual demonstrative* file, is indeed relevant to the behaviours that are rationalized by the attitudes that it contains (Recanati, forthcoming). What is important, for our purposes, is that this does not interfere with the Relationist construal of the file-framework. Though file-type is a representational feature, it is not one that individuates files: there can be, for example, distinct distinct perceptual demonstrative files about the same object (cf. Austin's (1990) 'two tubes' case). So file-type does not play the role of sense *vis-a-vis* coordination: sameness of file-type does not determine coordination.

A final way of making the point: we can see whether file-containment constitutes an intrinsic representational feature by examining what kinds of *permutations* on an attitude state are representationally significant. If file-containment is a representational feature, permutations that preserve the structure of co-filing but change which predicates are stored in which files could alter the representational features of an attitude state. But the file-framework doesn't appear to make use of meaningful permutations of this kind.<sup>29</sup>

Imagine for example that Smith has two distinct name-based files on Twain/Clemens containing the predicates 'is called 'Twain'' and 'is called 'Clemens'', respectively. She also has beliefs to the effect both that Twain is from Connecticut and to the effect that Clemens is from Connecticut, giving her two bodies of attitude, each about Twain/Clemens, which are not coordinated with one another. Let's call Smith's Twain file *A* and her Clemens file *B*. The file-theorist thinks that there is a representation of Clemens/Twain as a Connecticutian in both *A* and *B*. But would she say that the fact that one representation is in *A* and the other is in *B* is an intrinsic representational difference between the two representations of Twain as a Connecticutian? That is, abstracting away from the information in *A* and *B* (for example, the information concerning what the individual in question is called), do the individual identities of

---

<sup>28</sup> Thanks to a reviewer for prompting us to clarify here.

<sup>29</sup> Relationists have deployed this line of argument against Fregeanism (see, for example, (Pryor, 2016. pg. 334)). For an extended discussion of what it shows, see (Gray, manuscript).

*A* and *B* constitute representational differences between the two representations of Twain as a Connecticutian? We think not. Suppose we took all the predicates in *A* and put them in *B*, and *vice versa*. In making this switch, would we have thereby changed Smith's representational state? The file theorist should answer that we have not (it isn't clear, in fact, that this is even a coherent permutation for the file-theorist). After all, filing facts were meant to explain rational permission to infer (and act, and so forth) but the purely haecceitic switch does not change what Smith is rationally permitted to do.

Absent some argument--which file-theorists have not given--about how file *identity* makes a difference to rational explanation, the file-theorist should be a Relationist about coordination. We make this claim despite the official rhetoric of file-theorists to the effect that files, qua mental particulars, are *modes of presentation or senses*.<sup>30</sup> According to us, if files exist, they do not play the role of senses in a Fregean account of coordination: file-containment does not constitute an intrinsic representational feature that determines coordination. If files have an explanatory role to play, then, it must not be a semantic role but a role at another level of explanation.

## 6 The Metasemantics of Coordination

Supposing, then, that coordination is an irreducibly relational representational feature, is there any explanatory role left for files to play in relation to it? In this section we consider the metasemantics of coordination and argue that we can capture the basic metasemantic commitments of the file-theory without appealing to files.

Given CRR, we are committed to thinking of a subject's attitude state as consisting, in part, of a class of object-representations and coordination relations holding between some of them. If coordination is an equivalence relation (as we have been assuming), this structure is consistent with an explanatory role for files (we could think of each file as determining an equivalence class of coordinated representations). But if we are committed to object-representations anyway, why do we need the file? Why not go fully relational?

---

<sup>30</sup> See (Recanati, 2015, p.12, p.71) for just a couple of examples.

Astute readers will have noticed that an aspect of the file-picture has dropped out of the discussion in the previous two sections: ER relations. ER relations were supposed to play a constitutive role in relation to file-individuation, and by extension, trading on identity. It's natural to suppose, then, that if files have an explanatory role to play, it will involve the connection between files and ER relations. And it might be suspected that since ER relations are supposed to govern files *as a whole* -- rather than this or that representation contained in a file -- there will be no way to understand the role ER relations play without acknowledging the existence of files.

But this is not the case. We will show that the role that ER relations play in connection to trading on identity in the file-theory can be understood by thinking of ER relations as part of the metasemantic story for coordination. And this can be done without appeal to files.

Let's start by noting that once we see that coordination is an irreducibly relational representational feature, it becomes obvious that there is a metasemantic question about it. For any representational feature that we posit for an attitude state, we can ask: in virtue of what does the attitude state possess that representational feature? This question is just as sensibly raised for representational relations as for representational properties.

Our claim is that the file-theorist should hold that ER relations are part of the metasemantic story for coordination. Note that it is already a standard part of the file-approach that ER relations play a metasemantic role: they determine the *reference* of attitudes (Recanati, 2012, pg. 38). What is novel about our story is that they also play a metasemantic role in relation to coordination.<sup>31</sup>

---

<sup>31</sup> The general idea here--that is, the one expressed vaguely at the end of Section 2 of the paper--is not new. It is expressed in various ways in such works as (Evans, 1982), (Campbell, 2002), (Recanati, 2012), and (Dickie, 2015). But, it hasn't been clearly articulated. In particular, previous articulations do not make the distinction, which we take as crucial, between the rational status of a relational representational fact and the functional metasemantics of that fact. It should be mentioned that Dickie (2015) is clearest about the idea that the processes by which information is managed (what she calls, 'information-marshaling strategies') provide the grounding of the rational status of inferences that trade on identity. But our package of commitments differs from hers. She treats coordination as an 'activity' (2015, 86) and then claims that this activity must itself be justified (thereby motivating her to appeal to a story about practical justification). We claim instead that coordination facts are representational facts, and that ER relations play a metasemantic role with respect to those facts.

A preliminary note: we would like to make this claim without going in for any particular style of metasemantics (interpretationist, information-theoretic, teleological, *etc*). We think that whatever style of metasemantics you favour, you can accept the outlines of what we offer here. So we try to work at a level of generality that will find traction in different styles of metasemantics.

To that end, we will rely on (what we take to be) an unobjectionable metasemantic principle: the representational features of an attitude state are constrained by its functional properties. This is a weak claim. It only says that the functional structure of a subject's mental economy places constraints on the representational properties of her attitudes. It should be obvious why information-theoretic approaches to metasemantics endorse this principle (they hold that functional properties -- broadly construed so as to include covariational relations to worldly states -- determine representational properties). The same is true for teleological approaches (whose difference from information-theoretic approaches consists in their conception of the relevant kinds of functional properties). The situation is somewhat more complicated for interpretationist approaches; but we will suggest below that the kind of functional constraints we posit below will fall out of a plausible interpretation of the principle of charity.

We should also mention that we are not assuming that a full reduction of representational properties to non-representational properties is possible. We are only assuming that there are non-trivial connections between representational properties and functional properties.

We are interested in the metasemantics of a representational relation -- *viz.* coordination -- so we must focus on the way that distinct object representations are functionally *related*. We will show how it is possible to characterize the functional underpinnings of coordination in a way that draws on the resources of the file-theory, without actually appealing to files.<sup>32</sup> Our strategy will be to characterize a two-part functional role that underpins coordination. Roughly, we claim that two of *S*'s representations of the same object

---

<sup>32</sup> Clark (2017) also discusses the functional underpinning of coordination. He doesn't appear to think of this as a coordination as semantic, though, so the nature of the discussion is very different from the present one.

are coordinated only if *S* is disposed to trade on their identity and an ER relation plays a certain role (to be specified below) with respect to that disposition.

We are working with the background assumption that functional roles have two aspects: a ‘forward looking’ aspect and a ‘backward looking’ aspect (these are roughly analogous to elimination and introduction rules in natural deduction systems).<sup>33</sup> The backward-looking aspect specifies the characteristic way in which representations come to occupy that role. The forward-looking aspect specifies the characteristic way that representations occupying that functional role produce downstream effects. ER relations have a role in characterizing the backward-looking aspect of functional role that underpins coordination.

But let’s start with the forward-looking aspect. We have already seen that coordination is connected to a certain class of rational transitions: trading on identity. We propose, then, that part of the coordination functional role is the disposition to trade on identity; two representations of the same object are not coordinated unless their subject is disposed to trade on their identity under the appropriate circumstances.

This is vague. In particular, to spell it out we would need substantive characterizations of ‘trading on identity’ and of the relevant circumstances. We have inherited the vagueness of ‘trading on identity’ from the file picture (see discussion above). So we are content (for current purposes) to leave that imprecise. We will also not give a theory of ‘appropriate circumstances’. But the idea should be intuitive: our metasemantics will not count two representations as coordinated unless, if an occasion for trading on their identity arose, other things being equal, the subject would trade on their identity.

By an ‘occasion for trading on identity’, we simply mean: a cognitive situation in which trading on identity would suit some cognitive purpose for the agent, the agent has no other overriding cognitive purposes, the agent is aware of the relevance of the to-be-traded-upon attitudes to their cognitive purpose, *etc.*

---

<sup>33</sup> This sort of division is at least implicit in much theorizing about functional role. Our explicit inspiration for this framework comes from (Campbell, 2002) and (Ninan, 2016).



To give an artificial example, imagine a different version of our bird-watching scenario: I'm watching a bird, *b*, fly across the sky. At  $t_1$ , I judge that *b* has a yellow bill. At  $t_2$ , I judge that *b* has a graduated tail. At  $t_3$ , the question occurs to me: "I wonder if any birds have yellow bills and graduated tails?". I think to myself, "I just saw a bird with a yellow tail, and I just saw a bird with a graduated tail. But I don't think I'm in a position to answer this question". We take it that this would be strong evidence that, at  $t_3$ , the two beliefs about *b* were not coordinated.<sup>34</sup>

We can think of this constraint as an extension of the principle of charity. It would be deeply uncharitable to treat an agent as in a representational state that immediately licenses a simple inference if the agent is not disposed to make that inference when it would be in their cognitive interest to do so. That would be to treat the agent as rationally unintelligible.

Let's call this aspect of the coordination functional role, the trading on identity functional aspect (TIF).

TIF: object representations *a* and *b* stand in TIF if and only if, if an occasion for trading on their identity arose, other things being equal, the subject would trade on their identity.

We expect the idea that TIF is part of the metasemantics of coordination will be uncontroversial for file-theorists. Our claim is that the file-theorist should hold that this is only *half* of the functional underpinning of coordination. Recall that the file-theory connects the individuation of files – and thus the obtaining of coordination relations – to ER relations. We are getting rid of files, and so will reinterpret this connection at the level of the metasemantics of coordination. The backward-looking functional role of coordination is this: two representations of the same object are coordinated only if the fact that they stand in TIF is connected in the right way to the ER relations that the subject stands in.

A difficulty we face here is that the notion of ER relation is not given a precise definition in the file literature. We can glean an outline: an ER relation is a relation that a subject can stand in to an object; standing in that relation produces, under characteristic circumstances, doxastic states in the subject (paradigmatically beliefs, but perhaps also perceptual

---

<sup>34</sup> It would not be decisive evidence because the disposition, like all dispositions, could be present but blocked.

representations or some other kinds of accuracy-apt representations); across normal (or, perhaps, epistemically favourable) circumstances, the relation produces representations that are about the same object.<sup>35</sup>

Clearly this characterization is vacuous without a substantive account of normal circumstances for an ER relation. We're not aware of any general discussion of this in the file-literature. But if we look at the kinds of ER relations that file-theorists appeal to -- perceptual tracking, memory of a perceptual encounter, recognitional ability, competence with a proper name, knowledge of a descriptive condition that applies to the object -- we can imagine giving relatively substantive, though probably not fully reductive, characterizations of the relevant conditions. So we will assume that we have a tolerably clear understanding of ER relations.

Recall that we are trying to understand the role of ER relations in the file-framework without appealing to files. Our question is: what is it for two representations of the same object to be governed by the same ER relation if not to be contained in a file that is governed by that relation? As before, we will characterize a functional relation that two representations can stand in. The important idea here is to connect ER relations directly to TIF. We will think of the aspect of the coordination functional role contributed by ER relations (ERF) as follows:

ERF: object representations *a* and *b* stand in ERF if and only if there is an ER relation *R* such that if object representation *c* is a deliverance of *R* then, *ceteris paribus*, *c* stands in TIF to *a* and to *b*.

The idea is simple: representations *a* and *b* are governed by the same ER relation if and only if any object representations that are generated by that ER relation will stand in the trading on identity functional role to both *a* and *b*.<sup>36</sup> That is, the file-theorist should say that ER relations

---

<sup>35</sup> This is a place where the non-reductivity of the analysis shows: we presuppose facts about what the representations delivered by ER relations are *about* in our characterization of ER relations (and thus in our characterization of the functional constraint on coordination). As we said above, file-theorists have not offered a precise account of ER relations. If they have a non-semantic characterization of them, we could insert it here.

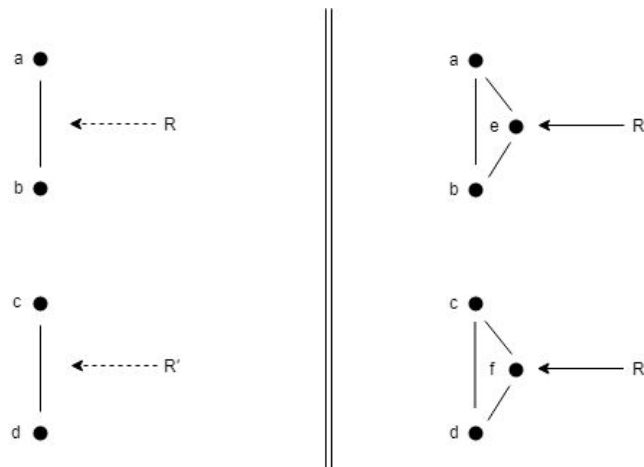
<sup>36</sup>It might be noted that ERF presupposes the diachronic individuation of object-representations: it characterizes the functional connections of an object representation before and after an ER relation outputs a new representation. We take this to be unproblematic. But it could be avoided, if desired, by adopting a stage theory

don't simply generate representations that are about the same objects, they generate representations that stand in TIF to each other.

We will motivate this idea shortly. But first, let's put the whole package together. Our claim is that the metasemantic functional constraint on coordination that can be extracted from the file-theory is the conjunction of TIF and ERF.

Metasemantics of Coordination:  $a$  and  $b$  are coordinated only if  $a$  and  $b$  stand in TIF and ERF.

The claim is that coordination is the upshot of the interaction between two kinds of functional organization. The diagram below represents the situation in which the functional relations between object-representations  $a, b, c, d$  are such as to ground coordination between  $a$  and  $b$  as well as coordination between  $c$  and  $d$ . The lines between nodes represent those nodes standing in TIF. For each pair of connected nodes, there is an ER relation ( $R$  for  $a$  and  $b$ ,  $R'$  for  $c$  and  $d$ ) that will deliver new object representations ( $e$  for  $a$  and  $b$ ,  $f$  for  $c$  and  $d$ ) that also stand in TIF to each antecedently connected representation.



Note that the way we've set things up, for two representations to be coordinated it's not required that each is actually the upshot of the same ER relation. The role of ER relations is

---

for object-representations -- as in (Prosser, 2019) -- and framing ERF in terms of suitably related sequences of object-representations.

dispositional: that two representations are governed by the same ER relation is a feature of how they are jointly situated with respect to dispositions to acquire new information. It is consistent with our claim that two representations are coordinated even though they were acquired through different ER relations (for example, perhaps one was acquired perceptually and the other was acquired testimonially)<sup>37</sup>. What we do demand, is that the way that they occupy TIF is dispositionally connected to an ER relation. In particular, we require that they are situated so that, *going forward*, if a new state stands in TIF to one of the representations, it stands in TIF to the other as well. This strikes us the most faithful reconstruction of the file-theorist's idea that files are governed by ER relations (if we wanted a more restricted picture -- one that requires that coordinated object representations have a certain etiology -- it would be straightforward to add this to our metasemantic story).

We have already motivated TIF as a metasemantic constraint. It remains, then, to motivate ERF. Consider a creature whose cognitive life is different from ours. Some of their attitude states stand in the TIF relation. That is, they are sometimes disposed to trade on the identity of various of their attitudes (or, if trading on identity presupposes coordination, they are disposed to do what looks like trading on identity). But that disposition is unrelated to any tracking ability. To employ the file-metaphor: suppose that the creature randomly sorts incoming information into files. Our claim is that none of the creature's attitudes would be coordinated. This is to say, remember, that they are not in a representational state that *rationalizes* trading on identity. How could we see the disposition to trade on identity as rationally relevant if it is not at least the typical downstream effect of a process that reliably delivered information about a single object? If the creature finds itself with the disposition to trade on the identity of two pieces of information derived from the same object, this will be a matter of pure chance. As such, it doesn't deserve any rational credentials.<sup>38</sup> The principle of charity has limits; and we reach its limits when conferring unearned representational features

---

<sup>37</sup> This allows for coordination relations to hold between representations to hold in a case like the one illustrated in Section 3.

<sup>38</sup> Goodman (manuscript) makes this point.

on an attitude threatens to sever the constitutive connections between representation, rationality, and non-accidental cognitive success.<sup>39</sup>

A final note: the discussion in this section has been proceeding in our own voice, but we should be slightly more careful. While we don't think this approach is implausible, we are primarily concerned to establish that this is what the *file-theorist* should say. We have offered a translation of the claim that files are synchronically individuated by their ER relations, into a claim about the metasemantics of coordination. Our translation captures the file-theorist's guiding idea that the rational credentials of trading on identity are grounded in cognitive abilities to track objects. And it does so without mentioning files.

## 7 Filing without Files

At this stage, we have illustrated how we can take a standard version of the file theory and re-write its explanatory claims without appeal to files. This means we can think of files as metaphors rather than theoretical posits. Mental files are merely a useful metaphor for talking about the way that two kinds of functional organization interact to ground irreducibly relational representational features of attitude states.

Note that our 'file free' account appeals to the same explanatory resources as the standard file theory does: *trading on identity, coordination, ER relations* etc.<sup>40</sup> Therefore our claim is not that the file-literature has the wrong explanatory resources. Rather, on our view, taking the file metaphor too literally has gotten in the way of seeing how those resources interact to explain what they are meant to explain. Once we clarify this, we see that files themselves don't play an explanatory role, and so don't earn a place in a canonical statement of the theory. We can countenance *mental filing* -- that is, the integrated dispositions associated with ER relations and trading on identity -- without positing any mental particulars that play the role of files.

---

<sup>39</sup> This connection is stressed in (Fodor, 1994) and, in a different way, in (Dickie, 2015).

<sup>40</sup> Note, it dispenses with appeal to the problematic containment metaphor but, in appealing to an *activity* of mental *filing* that plays a metasemantic role with respect to relational representational facts, it is also distinct from standard non-file-theoretic views of concepts that invoke neither of those things.

Before we comment on the significance of this work, a word about limitations. We've been focusing on the putative explanatory role that files play in relation to trading on identity. However there are other roles that files are said to play. To mention two, files are supposed to explain *continued belief* and are also sometimes said to explain certain facts about *interpersonal communication*.<sup>41</sup> So, in a sense, our work is unfinished. It is possible that these roles for files justify positing files as mental particulars, but we are skeptical. However, we will note that Relationist accounts of these phenomena have already been proposed.<sup>42</sup> And if they are on the right track, we could adopt them in our framework.

Returning, then, to the significance of this work. The fundamental point is simple: re-writing the file-theory without the file-metaphor brings out what the theory is really committed to. As we have said, readers are sometimes confused as to whether the file-approach is theory of content or a theory of the functional organization of the mind. Re-writing the theory without files lets us see that it is both, and allows us to understand their interaction: the file-theory offers a partially relationist theory of attitude content, as well as a theory of the functional grounding of relational representational facts.

A concrete benefit to eliminating files (and appealing only to *filing*) is that we can cleanly avoid the puzzles about containment. One worry, recall, was about relational belief: if I believe that *a* stands in *R* to *b*, does that correspond to the predicate '*x* stands in *R* to *b*' being contained in a file about *a* or the predicate '*a* stands in *R* to *x*' being contained in a file about *b*? On our approach, the puzzle evaporates. Believing that *a* stands in *R* to *b* simply corresponds to the presence of a belief that *a* stands in *R* to *b* in which the representation of *a* is coordinated with different further representations of *a* and the representation of *b* is coordinated with different further representations of *b*.<sup>43</sup>

---

<sup>41</sup>Perry (1980) introduces files to account for continued belief, and this is a theme of Recanati's discussion of cognitive dynamics in (2012) and (2016). Recanati deploys the file-framework in an account of interpersonal communication in (2012, chps, 14-16) and (2016, Part 2).

<sup>42</sup> See (Fine, 2007), and (Onofri, 2017) for Relationist accounts of interpersonal coordination. Prosser (2019) discusses both interpersonal and diachronic coordination. Strictly speaking, he offers a stage-theoretic account of files-qua-continuants. But, his general approach could be adopted without commitment to files (which he seems to view as metaphors, in any case).

<sup>43</sup> Pryor (2016, pg. 328) makes a similar point about his own graph-theoretic approach to propositional attitudes. Pryor's approach is broadly Relationist, and is one of the inspirations for the present project. We differ from Pryor, centrally, in trying to hold onto more of the explanatory framework of the file approach and, relatedly, by considering the metasemantics of coordination.

Another worry was about reciprocal containment: if we posit files, there is a sense in which a file about  $a$  contains beliefs about  $a$ ; there is also a sense in which my beliefs about  $a$  contain my  $a$ -file (either as vehicular constituent or as sense). Again, on our approach the puzzle evaporates: the representational features of a belief about  $a$  are partly constituted by the coordination between that representation of  $a$  and other representations of  $a$ . But there is nothing that resembles reciprocal containment.

The third worry was about trading on identity across belief and desire (or across attitude-type, more generally): containment cannot correspond to belief if trading on identity requires containment. Recall, it's not clear how we can explain in terms of containment (and therefore belief) the rationality of my looking to the night sky because I believe Hesperus is visible in it and desire to see Hesperus. But, if we replace containment with coordination, the problem disappears. Coordination, as a semantic relation, can hold between the content of a belief and the content of a desire. And our account of the metasemantics of coordination is also consistent with coordination across attitude-type. Our account holds that the coordination between my belief that Hesperus is visible and my desire to see Hesperus is grounded, on the one hand, in my disposition to trade on identity between them and, on the other, in the way that incoming information is related to such dispositions to trade on identity. In particular, our account requires (instead of containment) that there is an ER relation such that formation of further attitudes about Hesperus on the basis of it would stand in the TIF relation to my belief and to my desire.

We noted earlier that the containment puzzles had their source in the awkward relation between file-structure and belief in the file-framework. On the one hand, files play the role of *concepts* or *senses*, and thus have a quite general explanatory role in relation to trading on identity. On the other hand, file-structure is constitutively connected to belief, reflecting the privileged role that doxastic states are supposed to play in explaining trading on identity.

Our framework captures both of these ideas, but without the awkwardness induced by the metaphor of containment. Semantically, coordination is perfectly general: it can hold between object representations that are constituents of attitudes of any type.

Metasemantically as well, any object-representations can stand in the relevant functional role. But belief plays a privileged role in the metasemantic story by way of ERF. ER relations generate beliefs (or, at least, doxastic states). So our framework captures the idea that a subject's doxastic relation to an object -- as opposed to say, her conative relation, or suppositional relation to it -- has a privileged role in the explanation of her ability to trade on its identity in various attitudes.<sup>44</sup> Two object-representations only stand in ERF if there is an ER relation that will generate *beliefs* that will stand in TIF to each of them.

Another benefit of our approach is that re-writing the file-theory without files illuminates how one might argue against the approach or how one might propose to alter it.

One way to disagree with the file-approach would simply be to argue against Relationism. Perhaps there are no such things as irreducibly relational representational features, or perhaps coordination is not among them. If that's right, the file-approach would be in trouble. Another way to disagree with file-approach would be to question its functional underpinnings: either by questioning the idea of an ER relation, or by questioning the assumption that trading on identity is constitutively connected to it. It certainly isn't obvious that for any coordinated body of attitudes, there is a substantive epistemic link to the object they are about.<sup>45</sup>

More interestingly, our discussion brings out how the standard file-approach can be seen as part of a family of related views. And once we've seen it that way, there may be pressing questions about what speaks for the standard version over the alternatives. We'll mention one important example here.

Looking back at ERF, note that it existentially generalizes over ER relations: two representations are coordinated only if there is *some* ER relation whose deliverances stand in TIF to both of them. Note that it doesn't fall out of this that coordination is an equivalence

---

<sup>44</sup> This is another way in which our view differs from existing, non-file theoretic accounts of concepts.

<sup>45</sup> An obvious worry here is that there are non-referring files. In general, we have been assuming that coordination implies coreference (and thus implies reference). This is a vexed issue (see (Lawlor, 2010), (Fine, 2010a)). If we want coordination to be consistent with failure of reference, the file-approach would need to develop a conception of a (pseudo-)ER relation that is consistent with the non-existence of one of the relata. Given such a conception, our metasemantic story can be told as before.



relation.<sup>46</sup> If it is possible for a representation to be governed by more than one ER relation, this constraint is consistent with an intransitive picture of coordination: *a* might share one of its ER relations with *b*, and *b* might share one of its ER relations with *c*, despite *a* not sharing any ER relation with *c*. If this is a possible distribution of ER relations, and we want the transitivity of coordination to fall out of the functional constraint, we would have to re-write ERF to reflect that coordinated representations must share *all* of their ER relations. (The analogous point can be made about TIF).

That coordination is intransitive is not a merely idle possibility. Consider an example sentence from Pinillos (p 314): “We were debating whether to investigate both Hesperus and Phosphorus; but when we got evidence of their true identity, we immediately sent probes there.” According to Pinillos, “Hesperus” and “Phosphorus” are each coordinated with “there”, but are not coordinated with each other. If that’s correct, we might think that what is going on with the attitudes of someone who sincerely asserts this sentence is that the object-representation expressed by “Hesperus” is governed by one ER relation, the object-representation governed by “Phosphorus” is governed by another, while the object-representation expressed by “there” is governed by both of those ER relations.

We don’t want to suggest that considerations of this kind are decisive. The point is rather to illustrate that, if we take the file-metaphor too seriously, moving to an intransitive notion of coordination requires completely abandoning the theory (because coordination-relations would no longer generate equivalence-classes of representations and so could not correspond to file-structure). But now that we have re-written the theory without mention of files, it looks like the theory can accommodate intransitive coordination. We may decide not to accommodate intransitive coordination, given our understanding of the theoretical role of coordination or of its metasemantics, but we shouldn’t forestall the possibility of doing so simply because we are tied to the file-metaphor.

## References

---

<sup>46</sup> There is a more basic reason it doesn’t fall out of it: it only gives a necessary condition on coordination. The point here is that this necessary condition doesn’t impose a transitivity constraint on coordination in relation to ER relations.

- Austin, D. F. 1990. What's the Meaning of "This"? A Puzzle about Demonstrative Belief. Cornell University Press.
- Campbell, J. 1987. Is Sense Transparent? *Proceedings of the Aristotelian Society*, 88, 273-292.
- Campbell, J. 2002. *Reference and Consciousness*. Clarendon.
- Clarke, H., 2018. Frege Puzzles and Mental Files. *Australasian Journal of Philosophy*, 96(2). pp.351-366.
- Dickie, I. 2015. *Fixing Reference*. Oxford University Press.
- Donnellan, K. 1978. Speaker Reference, Descriptions and Anaphora. *Syntax and Semantics* 9: 47–68.
- Evans, G. 1982. *The Varieties of Reference*, edited by J. McDowell. Oxford University Press.
- Evans, G. 1985. The causal theory of names. In *Collected papers*, 1–24. Oxford: Oxford University Press
- Fine, K. 2007. *Semantic Relationism*. Wiley-Blackwell.
- Fine, K. 2010a. Reply to Lawlor's "Varieties of Coreference". *Philosophy and Phenomenological Research*, 81(2), 496–501.
- Fine, K. 2010b. Semantic necessity. In B. Hale & A. Hoffman (Eds.), *Modality*, pp. 65–80. Oxford
- Fodor, J. A. 1994. *The Elm and the Expert: Mentalese and its Semantics*. MIT Press.
- Green, E. J., & Quilty-Dunn, J. (2017). "What Is an Object File?" *The British Journal for Philosophy of Science*. DOI: axx055.
- Grice, H.P. 1969. Vacuous names. In *Words and objections*, ed. Hintikka, D.D.J., 118–145. Dordrecht: Reidel.
- Goodman, R. 2016a. Against the mental files conception of singular thought. *Review of Philosophy and Psychology*, 7(2), pp.437-461.
- Goodman, R. 2016b. Cognitivism, significance and singular thought. *The Philosophical Quarterly*, 66(263), pp.236-260.
- Goodman, R. Manuscript. Singular Thought and Trading on Identity.
- Goodman, R. & Genone, J. 2020. Mental files and singular thought: Introduction, in *Mental Files and Singular Thought*, edited by R Goodman, J. Genone and N. Kroll. Oxford University Press, pp.1-17.
- Gray, A. 2017. Relational Approaches to Frege's Puzzle. *Philosophy Compass*, 12(10):e12429.
- Gray, A. Manuscript. Minimal Fregeanism.
- Heck, R. K.. 2012. Solving Frege's Puzzle. *Journal of Philosophy*, 109, 132–174. (Originally published under the name "Richard G. Heck, Jr.").
- Jeshion, R. 2002. Acquaintanceless de re belief, in *Meaning and Truth: Investigations in Philosophical Semantics*, edited by J. Campbell, M. O'Rourke and D. Shier. p.53-78. New York: Seven Bridges Press.
- Jeshion, R. 2009. The significance of names. *Mind & Language*, 24(4), pp.370-403.
- Jeshion, R. 2010. Semantic Instrumentalism, and Cognitivism. *New essays on singular thought*, p.105. Oxford University Press
- Lawlor, K. 2010. Varieties of coreference. *Philosophy and Phenomenological Research*, 81(2), 485–495.
- Murez, M., Smortchkova, J. & Strickland, B. 2020. The mental files theory of singular thought: a psychological perspective, in *Mental Files and Singular Thought*, edited by R. Goodman, J.Genone and N. Kroll. Oxford University Press, pp. 107-142.
- Ninan, D. 2015. On Recanati's Mental Files. *Inquiry: An Interdisciplinary Journal of Philosophy*, 58(4),

- Ninan, D. 2016. What is the problem of De Se attitudes? Pages 86–120 of: M. Garcia-Carpintero, S. Torre (ed), *About Oneself: De Se Attitudes and Communication*. Oxford University Press.
- Onofri, A. 2015. Mental files and rational inferences. *Inquiry: An Interdisciplinary Journal of Philosophy*, 58(4), 378-92
- Onofri, A. 2017. The Publicity of Thought. *The Philosophical Quarterly*, 68(272), 521–541.
- Papineau, D. 2013. Comments on Francois Recanati's *Mental Files*: Doubts about indexicality. *Disputatio*. 5(36), 159-75.
- Perry, J. 1980. A problem about continued belief. *Pacific Philosophical Quarterly*, 61(4): 317–322.
- Pinillos, N. Á. 2011. Coreference and Meaning. *Philosophical Studies*, 154(2), 301–324.
- Prosser, S. 2019. The Metaphysics of Mental Files. *Philosophy and Phenomenological Research*. doi:10.1111/phpr.12577
- Pryor, J. 2016. Mental Graphs. *Review of Philosophy and Psychology*, 7(2), 309–341.
- Recanati, F. 2012. *Mental Files*. Oxford University Press.
- Recanati, F. 2015. Cognitive Dynamics: A New Look at an Old Problem. *Reference and Representation in Language and Thought*, edited by K. Korta and M. Ponte.
- Recanati, F. 2016. *Files in Flux*. Oxford University Press.
- Recanati, F. 2020. Coreference de jure. *Mental Files and Singular Thought*, edited by R. Goodman, J. Genone and N. Kroll. Oxford University Press, pp. 162-186.
- Recanati, F. forthcoming. Reference and singular thought. *Routledge Handbook of Linguistic Reference*, edited by H. Geisson and S. Biggs.
- Taschek, W. 1995. Belief, Substitution, and Logical Structure. *Noûs*, 29(1), 71–95.
- Taylor, K. A. 2003. *Reference and the Rational Mind*. CSLI Publications.
- Woodfield, A. 1991. Conceptions. *Mind*, 100:547-72.