Mental Files:
Replies to my Critics

François Recanati
Institut Jean Nicod, Ecole Normale Supérieure

1 The primacy of acquaintance

Coliva and Belleri write:

> '[An] aspect of Recanati's proposal that is not entirely clear is its precise scope. For we are told that mental files are mental indexicals which depend, for their existence, on there being an epistemically rewarding relation, in the form of acquaintance, between a subject and the object the file is a file of. On the face of it, however, this would entail that mental files are quite limited, for we do not seem to be acquainted with a lot of entities that we are nevertheless able to think about. We are not acquainted with non-existing and fictional entities; nor are we acquainted with past or future entities, let alone with abstract ones, like numbers or logical entities.' (Coliva and Belleri: 109-10)

Coliva's and Belleri's observation raises a dilemma for my view: either its scope is very limited, as they suggest — it only accounts for a small class of mental files based on acquaintance; or — if the account is taken to apply to all mental files, including those that are purely descriptive because there is no way in which we could be acquainted with the object they are about — then it is totally implausible.

Clearly, I offer my theory as a *general* account of singular thought; so its scope should *not* be limited to a small sub-class. This means that I have to embrace the second horn of the dilemma and face the implausibility criticism. As Keith Hall puts it,

> 'Instances of mental files (…) whose function is ostensibly not characterizable in terms of acquaintance are neither exotic nor rare: on the contrary, they pervade our cognitive lives. At the very least, the proliferation of such problem-cases should make us question the motivation for taking the acquaintance-based function of files to be primary.' (Hall: 129)

This worry is especially pressing since there is no shortage of theorists who reject the acquaintance constraint on singular thought. All the mental files 'whose function is ostensibly not characterizable in terms of acquaintance' lend prima facie support to their approach, as against mine. As a result, 'If Recanati is to maintain that the acquaintance-based function of his mental files is somehow primary, he needs to offer more by way of argument for that claim' (Hall: 129).

Let me start by rehearsing what I actually say in the book, in order to make clear what I mean when I talk of the primacy of acquaintance. I say that mental files are governed by an acquaintance norm. Acquaintance is understood liberally, so that testimony counts as (mediated) acquaintance. Tokens of a given file carry the presupposition that the norm governing the file is satisfied: they carry an acquaintance presupposition. The acquaintance presupposition generates what (in a two-dimensionalist framework) we may refer to as the primary content of the file token, distinct both from its referential or secondary content, and also from its informational content (the information in the file).

When the norm/presupposition is violated, the file fails to express a secondary content — it fails to refer — but it still carries its primary content. So I can accept that, in such cases, the subject is still 'thinking a singular thought', in one sense of that phrase: a singular vehicle is tokened, which carries the primary content it inherits from the type of which it is a token. This accounts for the cases of de facto non-acquaintance, e.g. for the cases in which the subject is thinking about a nonexistent object such as Vulcan.

There is another type of case, in which there is no 'failure' properly speaking. There is no failure because the subject is not in the business of thinking about objects s/he is or might be acquainted with. Thinking about fictional characters like Emma Bovary is a case like that; thinking about the average American male, and calling him 'Bob', may be another case of the same sort. All such cases are characterized by the following two features: (i) there is no attempt to satisfy the norm, yet (ii) the norm remains operative. The norm remains operative because it is 'exploited', to use Grice's phrase. (According to Grice, the maxims of conversation are exploited when one flouts them blatantly in order to convey an implicature.) What the subject is doing with the file in the relevant cases is not governed

by the norm — so the subject can't be criticized for violating the norm — but doing what he or she does *still presupposes the norm*. The subject could not do what s/he is doing with the file if the norm wasn't in place. I will henceforth use 'exploitation' as a technical term to refer to the cases in which these conditions are satisfied.

The most general exploitation mechanism is *simulation*: one does *as if* one were using the file normally (i.e. according to the norm) to refer to an object of acquaintance. According to a simple and powerful story, this is how fiction works (Lewis 1978, Walton 1990, Kripke 2013): it involves the simulation of (perceptual or testimonial) acquaintance. It is pretty clear that simulation is parasitic on what is simulated, so condition (ii) is satisfied: doing as if one were following the norm presupposes the existence of the norm.

What should we say about the (secondary) content of the file in exploitation cases? Does the Emma Bovary file activated in the reader of *Madame Bovary* refer? The answer ought to be negative: there is simulated reference, but no actual reference. To be sure, there is a clear difference with the Vulcan case. In the Vulcan case the subject (Leverrier) makes a mistake and fails to think something true or false. In the Emma Bovary case, the reader makes no mistake and cannot be criticized. But the reason for that is that his or her thought occurs in a special, fictional mode. It does not matter that the files deployed by the reader of a fictional text fail to refer. As Frege suggested, in fiction the step from sense to reference is not taken. In my framework, that means that the primary content is sufficient — one does not care about the secondary content, so it does not matter if there is no secondary content for the fictional thought. Still, according to Frege, what the subject is thinking is neither true nor false.

What about descriptive names? Here also the thinker (the person who uses or understands the descriptive name) is not acquainted with the object — the object is given only by description; and here also, the thinker cannot be criticized for not being acquainted with the object. This suggests that, perhaps, there is a form of simulation in this type of case as well. That's the sort of story I floated in *Direct Reference*. In using descriptive names, arguably, we do *as if* the object was given (and use a singular vehicle to think about it). Why do we do that? In many cases (those I focussed on in *Direct Reference*) because we anticipate a forthcoming epistemic state in which we will be ac-

quainted with the reference. In other cases, there may be no such anticipation, but using a singular vehicle may yield worthwile cognitive dividends, by making thinking and reasoning easier (especially when counterfactual thinking/reasoning is involved).

If we opt for a simulationist approach to descriptive names, we can maintain Evans's hard line regarding their secondary or referential content. The strong acquaintance view, as I call Evans's position, says that the singular thought we attempt to think (about the object known by description) is not available for us to think. (This is reminiscent of what Russell says about 'Bismarck'.) The thought fails to have a singular content, though phenomenologically it feels as if it had a singular content. But tokening the thought vehicle in the simulative mode brings cognitive dividends which compensate for the local 'lapse into fiction' induced by descriptive names.

The strong acquaintance view is not the only possible take on descriptive names, even within a simulationist framework. (As I said already, I take the 'anticipated acquaintance' account of descriptive names put forward in *Direct Reference* to be a form of simulationism. To anticipate an acquaintance situation is to do as if it already obtained.) We can also liberalize the acquaintance constraint and say that when we *correctly* anticipate some forthcoming acquaintance relation to the denotation of the description, then we are in a position to think a genuine singular thought about it. This view (which Hall calls 'Loophole') is put forward very tentatively in *Mental Files*. I don't endorse it in the book — I mention it as an intermediate position (worth investigating) between the strong acquaintance view and liberalism. Hall and Coliva & Belleri don't like this intermediate view, which seems to posit a weird kind of backward causation. Moreover, Coliva and Belleri point out that it does not fit my picture very well, because the relation to the future referent is not epistemically rewarding. In any case, I don't really care whether Loophole is true or not. I can live with the strong acquaintance view, as I can live with the liberal view put forward by advocates of semantic instrumentalism.

Radical forms of semantic instrumentalism, such as that put forward by David Kaplan (1989a, 1989b), provide an alternative to simulation approaches. According to radical instrumentalism, using singular vehicles when the referent is known only by description

yields a 'broadening of the horizon of thought': a singular thought content which, without language, would *not* be available to the subject (because the acquaintance constraint is not satisfied) *becomes* available when we help ourselves to the linguistic apparatus of direct reference. On this view the thinker actually manages to refer singularly, so what may have started as a form of simulation has led to a form of 'cognitive restructuring': the acquaintance constraint is no longer operative and we can think singularly about objects through essentially linguistic means. This is very different from the strong acquaintance view, because the thought is now taken to have a full blown secondary content which is singular. If we follow this route we embrace a strong form of liberalism. We reach the conclusion that 'any act of descriptive reference-fixing would allow subjects to entertain singular contents without acquaintance' (Hall: 126).

Hall thinks this liberal position is not open to me:

> 'Although that conclusion would be welcome to me, this is precisely the sort of liberal view Recanati wishes to avoid. For this reason, I do not think that Recanati's semi-liberal position here is sustainable.' (Hall: 126)

By 'Recanati's semi-liberal position' Hall refers to the view that, although the acquaintance constraint holds, it is satisfied when we correctly anticipate a future acquaintance relation (Loophole). Now, as I said already, I don't really care about Loophole's fate. And I have no quarrel with semantic instrumentalism. I *can* accept the liberal view that some files are not governed by an acquaintance constraint. Appareances notwithstanding, the liberal view is compatible with my account for, in addition to mechanisms of exploitation, which presuppose the acquaintance norm, I also make room for *derived functions* for singular vehicles. When a singular vehicle acquires a derived function, the norm corresponding to the initial function no longer holds. So I am open to the instrumentalist suggestion that language 'broadens the horizons of thought' and makes it possible to use our mental files to do something other than what it is the primary function of the files to do.

'Primary' here must be understood in a new, evolutionary sense. The idea is this. We start with referential devices which are already in place in perception (as well as in infant cognition). These devices are encoded in language, but language makes it possible to do things

with them that we can't do in perception (or that infants can't do). For example, in language referential indices can be bound (through quantifiers which manipulate assignment functions instead of letting the context assign referents to the indices), and they can also be used freely, without external anchoring, as in E-type anaphora. I accept that there are these things in language (and in thought informed by language). So I am open to the idea that there are *purely descriptive files*, as people as diverse as Jeshion, Dickie and Cumming have recently claimed. That's compatible with my framework if one treats such descriptive files as a late achievement made possible by language.

Hall objects to my view that

'even if the acquaintance-based function of files is somehow evolutionarily basic, I see no compelling reason to think that the functions of files should not have evolved from their original acquaintance-based function.' (Hall: 130, note)

But that's exactly what I am saying. The cases which are the most striking counterexamples to the acquaintance view, namely the cases in which there is *not even exploitation of the acquaintance norm* but, rather, a different use of the file altogether, are cases in which the file has evolved derived functions, distinct from the evolutionary basic function of storing information gained through acquaintance. I accept that there are such cases, but I maintain that acquaintance has got some sort of primacy. I am making an empirical hypothesis: that the object tracking system which exists in perception is used throughout cognition — even in high-level cognition, e.g. in thought about abstract objects. The units of the system — the files — acquire new functions when they are recruited in this way to do extra jobs in thought. The hypothesis is that language plays a crucial role in making these new functions available.

To sum up, there are three types of case to consider, and two senses in which the acquaintance function of files is primary. The three types of case are: (i) the prototypical cases in which a file is used to track an object one is acquainted with and to store information gained through acquaintance with it; (ii) the exploitation cases in which e.g. one does *as if* the file was used to track an object of acquaintance and to store information about it (even though it's not); and (iii) the cases in which the acquaintance norm is no longer operative because the files have evolved derived functions. The two

senses of 'primary' are: (a) in exploitation cases the acquaintance norm which governs the type is presupposed, even though the thinker is not following the norm but engaging in simulation. Simulation is asymetrically dependent on what is simulated, and that is what establishes the primacy of the prototypical acquaintance cases over the simulation cases. (b) The derived functions correspond to new jobs assigned to the files, but the referential function of the prototypical files is primary because *the files are what they are in virtue of belonging to the object tracking system* (where they have a referential function), and the new roles correspond to uses of *these files* to do new things thanks to, inter alia, the interaction with the linguistic system.

2 Vehicles, modes of presentation, and primary content

Some of my critics are puzzled by my oscillation between talk of files as quasi-syntactic vehicles and talk of files as modes of presentation:

> 'There seem to be (…) ambiguities in the way mental files are presented. On the one hand, we are told that they are singular Fregean senses, that determine the referents they stand for. On the other, we are told that they are similar to Fodor's terms in the language of thought. However, Fodor's concepts are only syntactically different and do not contain any semantically relevant material apart from their referent, nor is the latter determined by sense.' (Coliva and Belleri: 109)

> 'Mental files are the vehicles of singular thoughts. They belong to the system of mental representations. They are the mental counterparts to singular terms, and they *refer*, or are *supposed to refer*. (…) At the same time, they are modes of presentation, which is to say 'senses' (257), and *individual concepts*, i.e. thought constituents (64). For someone used to thinking about language, this is puzzling.' (Pagin: 136)

But there is no puzzle, really. Modes of presentation are whatever plays the mode-of-presentation-role (as defined through what Schiffer refers to as 'Frege's Constraint').[1] Beyond that, it is an open issue what exactly modes of presentation are — what plays the role. My claim in the book is that what plays the mode of presentation role is the vehicle (the file). By this I mean that it is possible for a rational subject to take different attitudes towards (what is in fact) the same object if the subject has two numerically distinct files about that ob-

---

[1] See Schiffer (1978: 180).

ject. The files don't have to differ semantically: they may be of the same type, have the same referent, and contain the same information (or misinformation).

Pagin describes a Fregean semantics for mental files as based on the following principle:

> (DIF) Any difference in cognitive significance between two mental files can be explained by their semantic difference.

Pagin thinks I accept the principle, but I don't. I reject (DIF). Pagin himself, in the paper, shows that a mental file theorist is not in a position to embrace (DIF). He gives the following example:

> 'Suppose X takes herself to see two moths flying around in her kitchen. She opens a file for each, alpha and beta, thinking of them as 'A' and 'B', respectively. She takes herself to see now A, now B. The acquaintance relations are indeed different in case there *are* two moths, one causing the opening of alpha, the other the opening of beta. But in case the subject in fact is mistaken, and there is only *one* moth causing the opening of both files, there does not seem to be any difference between the acquaintance relations of alpha and beta. (…)

> Such a situation is certainly possible (and so are others, essentially like it). Is there any semantic difference between the files alpha and beta? A description theorist can certainly say that the terms 'A' and 'B' differ in sense: the one can have the sense of the description 'the moth I saw first' and the other the same sense as 'the moth I saw second', even if she coined the terms only after taking herself to have seen both and did not then remember the original sightings. (…) But this option is not open to the mental file theorist, since for both files there is an acquaintance relation to an object, in fact the same object, and there does not seem to be any difference between these relations *except* the distinctness of the mental file relata.' (Pagin: 140-1)

In the case described by Pagin, there is no semantic difference between the two files, which are exactly alike. Yet in that situation the subject can contemplate the possibility that alpha is *F* while beta is not *F*. What this suggests is that two files can differ in cognitive significance *solo numero*. So we should reject (DIF). That is what I mean when I say that modes of presentation are vehicles.

This does not mean that *only* the numerical identity of the file *qua* mental particular matters to cognitive significance; that would be absurd. My claim is weaker: nothing less fine-grained than the vehicle, and in particular no semantically defined equivalence class,

will be able to play the mode of presentation role in its entirety. We need mental particulars (as Perry often emphasized). But the vehicles have semantic properties, and among their semantic properties some contribute in a non-negligible way to cognitive significance. In particular, the type of a file carries a presupposition to the effect that a certain type of ER relation is in place. The presuppositions carried by a file in virtue of its type constitute its *primary content*.

Coliva and Belleri complain that, if files are vehicles, the theory fails to account semantically for what is common to singular thoughts when the acquaintance constraint is satisfied and when it isn't. What is common, in my framework, is only a singular vehicle. In the normal case, when the acquaintance constraint is satisfied, the vehicle has a singular content, but when the acquaintance constraint is not satisfied, a singular vehicle is tokened but it does not have a proper singular content. This is unsatisfactory, Coliva and Belleri say:

> 'Recanati describes subjects who entertain such files as thinking singular *vehicles* and not singular contents. To entertain a singular vehicle, he says, is to token a mental file which is not created on the basis of an acquaintance relation (either one that actually obtains, or one which is expected to actually obtain) (166-169). Singular vehicles however, are merely *taken* to provide singular reference by those who entertain them (if, e.g., they are mistaken about the existence of their referent — think of a child who believes in Santa Claus); at best, they are *treated* as providing singular reference (we may imagine a cautious scientist, who is not sure about the existence of the entity she is naming). In each case, theirs is only an *appearance* of singularity and it is not clear how one could go from an appearance of singular thought to singular thought proper, in any interesting semantic sense. So if entertaining a singular vehicle comes down to entertaining a *seemingly singular thought* (which is really not a singular thought, in any interesting semantic sense), we do not see how this notion could be of help.' (Coliva and Belleri: 110)

To sum up their argument: in a framework which takes the files to be syntactic entities, there will be no notion of singular thought characterizable 'in any interesting semantic sense'.

But I deny that. There is an interesting semantic sense in which to entertain a singular vehicle is to think a singular thought. The singular vehicle has a primary content, which it retains even when there is failure of reference (hence no secondary content). That is sufficient to ease Coliva's and Belleri's worry. The primary content corresponds to the fact that (even if the acquaintance constraint is

not satisfied) the subject entertaining the file is supposed to stand in the right ER relation to the object s/he purports to think about. The primary content has a normative character and survives when, de facto, the norm is not complied with.

Pagin objects to the idea that files have primary content in that sense:

> 'The concept of a mental file, like the concept of a *gene* in genetics, is in itself functional. Unlike linguistic expressions, we don't observe mental files or tokenings of them. We cannot demonstrate them. What we know about mental files, we know on the basis of the *theory* of mental files. 'Mental file' is a theoretical term in a theory like Recanati's, and the interpreted theory formulation it occurs in gives us its functional role, which is our concept of a mental file. We can then go on to ask whether there exist mental entities that fill this role, and whether there is a unique collection of entities that do. (…)

> In classical functional role theory (e.g. Loar 1986), the functional roles are causal. This means that we can investigate the causal pattern in e.g. processes of the brain, to find out what brain states fulfill certain functional roles. But when the functional role is *normative*, this cannot be done. We cannot, as a means of identifying mental files, get hold of a brain state, or a mental state, and ask whether that state is *required* to have an acquaintance relation to an object. A brain state, or a mental state, can at most be subject to such a requirement once it has been identified as a mental file by an independent criterion. For instance, if mental files were all causally related to external objects by some acquaintance relation, we could use that general fact to identify mental files. But this is not the case, on Recanati's account. It is only part of their function to be *required* to stand in an acquaintance relation. But that means conflating a consequence of being a mental file with what is needed for identifying them in the first place. That is why I find the idea of the primary content of mental files problematic.' (Pagin: 143-4)

Pagin says the problematic step 'comes when an *additional* functional role or normative requirement is included in the basic functional role itself' (143). But if this is the problem, then let's not include the normative requirement in the basic functional role. Let's start with the set of basic cases, i.e. the cases in which all goes well (the acquaintance relations are in place, etc.), and characterize the basic functional role of mental files at that level, as Pagin recommends. The assumption is that there are things that play that role, and that we can find out what they are by looking. Suppose we have identified them. Then a further assumption of the theory is that it will be possible to find *other things* with the following three properties:

(i)   They are of *the same type* (nonsemantically characterized) as the basic files.

(ii) They don't play the basic functional role because the acquaintance relations aren't in place.

(iii) Still, they have the same cognitive significance (by standard tests) as those involved in the basic cases.

The theory accounts for the putative observation in (iii) by saying that the type has acquired a primary content which derives from the function which the tokens play in the basic cases and is inherited by the tokens in the non-basic cases.

The distinction between primary and secondary content is relevant also to an issue Coliva and Belleri raise about transparency:

> 'One further feature of the theory which is not entirely clear is the extent to which one's singular thoughts are transparent. Recanati disagrees with both Boghossian's and Burge's different takes on the issue of the compatibility between externalism and self-knowledge. He claims that in the following kind of inference, taking place after a slow switch between worldly mental files and their counterparts on twin Earth,
>
> (1)   Jo once loved playing in the water
> (2)   Jo does not like playing in the water now
> (3)   Jo has changed
>
> it is not the case, *contra* Boghossian, that 'water' in (1) and (2) respectively refers to water and twater; nor is it the case, *contra* Burge, that the reference of 'water' in (2) is water, like in (1), because the reasoning initiated in (1) requires the reference of 'water' to remain stable. Rather, the reference of 'water' is confused in both cases, so it is neither water nor twater and therefore (1) and (2) are neither true nor false. Yet, according to Recanati, his account preserves transparency. (…)
>
> Even if one grants Recanati the idea of confused reference in (1) and (2), this would actually entail that while the subject may be thinking of thinking a (t)water-thought in each of the premises, he would not. So, it remains unclear how the proposed solution would actually allow to compatibilize externalism and the transparency of senses, for the content of one's thoughts would still be unknown to the subject.' (Coliva and Belleri: 111-2)

But the distinction between levels of content is crucial here. Reference is opaque, so secondary content is opaque as well. The subject does not know whether s/he is thinking about water, about twater,

or about nothing at all. What the subject has transparent access to is the mode of presentation, not the reference. Now in (1) and (2) the same mode of presentation — the same mental file 'water' — occurs twice, so the subject knows that these two occurrences co-refer if they refer at all. There are three options compatible with what the speaker knows in virtue of the transparency of modes of presentation: (a) the two occurrences both refer to water; (b) they both refer to twater; (c) they both fail to refer because the file's presupposition is violated in the slow-switch situation. Burge goes for option (a) (or option (b) if we reverse the order of the premises). I raise difficulties for his account, and advocate option (c), also compatible with the transparency of modes of presentation. This ensures (contra Boghossian) that a minimal degree of self-knowledge is compatible with externalism, but it has never been my intention to claim that externalism was compatible with a stronger, Cartesian form of self-knowledge corresponding to the transparency of secondary content.

3 Are mental files indexical?

There was a time when people took indexicality to be a property of language, and of language exclusively. Thought itself, they believed, *could not* be indexical (any more than it could be ambiguous).[2] Then came the discovery of 'the essential indexical'. Castañeda, Prior, Geach, Perry, Lewis and others put forward examples in which removing the indexicals from a sentence changes the nature of the thought that is expressed. This seems to establish that the thought expressed by uttering an indexical sentence is itself indexical, in some cases at least. Since then, the study of indexical thought has flourished. The nature of indexical thought has become a major topic in the philosophy of language and mind.

   Has the pendulum swung too far in the other direction? In his contribution to this symposium, Papineau suggests that it has:

   'There is a surprisingly widespread tendency to infer, from the use of indexical words to express some thought, that the thought expressed

---

[2] The reasoning behind that conclusion was fallacious — it begged the question against theories of thought that take the vehicle of representation to be constitutive of thought (syntactic theories, as we might call them).

must be similarly indexically structured. But it does not take much reflection on cases to show that this inference is generally invalid.' (Papineau: 171-2)

Papineau actually makes two claims against the 'mental indexicality' trend. One is that language and thought work differently — there is less indexicality in thought than language. The other claim is that the referential units in thought are *mental files coarsely individuated*. Such mental files are more like names than like indexicals. Papineau rejects my appeal to fine-grained mental files based on specific contextual relations to objects, on the grounds that 'personal-level conceptual thought has no good use for [them]' (9). As we shall see, he thinks that as soon as we encounter an object we open a *potentially lasting* mental file for it, abstracting away from our current ER relations to it. Now there is a type of file with the required characteristics in my system: 'encyclopedia entries' (as I call them) are lasting files, and they abstract from particular ER relations. Papineau likes them, but he does not think we need (in addition) the fine-grained files based on particular relations, such as demonstrative files.

Papineau makes a couple of sub-points pertaining to his first claim. First, he says that encyclopedia entries (which I treat as abstractly indexical because they rest on contextual relations to objects without specifying/requiring any particular ER-relation) are not indexical in the sense in which in language some words are indexical. He writes:

'The only good way to fit the personal encyclopaedic *Obama* files into the standard indexical type-token structure would be to view them as tokens of the type *encyclopaedic file* (or perhaps *encyclopaedic person file*). This type would have no reference of its own, and each of its tokens (such as an *Obama* file, or a *my first teacher* file, or a *that-woman-down-the-road* file …) would then have its referent fixed as that thing (person) in the relevant thinker's environment to which the relevant file bears some epistemically rewarding relation.

This would work all right, but it would be strange, and certainly wouldn't line up with any indexical constructions present in natural languages. We certainly don't have some type *word* the tokens of which refer variously to Obama, my first teacher, that woman down the road, … , depending on the context in which those tokens are uttered.' (Papineau: 163-4)

I admit that encyclopedic files are not indexical in the same sense in which demonstratives or the standard indexicals are. Whether or not we want to call them indexical is not a substantive issue (as Papineau notes). I call them indexical because the mode of reference determination is relational and corresponds to a feature of the type. The difference with standard indexical files in the narrow sense is that the relation is only specified abstractly.

Is Papineau right that this abstract form of indexicality does not line up with any indexical construction present in natural language ? I think that depends upon one's semantic theory. I myself have put forward a semantic analysis of proper names which treats them as indexical. (This is like Pelczar's analysis, which Papineau mentions in a footnote.) As Fiengo and May 2006 rightly stressed, a proper name *type* — what Kaplan 1990 calls a generic name — does not refer, but has (typically multiple) bearers, in virtue of various naming-conventions involving the name. What refers is a use of the name in a sentence. What a use of the name refers to is some object *x*, assigned in context (in virtue of speaker's intentions and related factors), *provided* the object in question satisfies the linguistic presupposition carried by the use of the name. What is presupposed is that *x*, the referent, is a bearer of the name.[3]

Papineau, as you recall, objects that

'We certainly don't have some type *word* the tokens of which refer variously to Obama, my first teacher, that woman down the road, ..., depending on the context in which those tokens are uttered.' (Papineau: 164)

But proper names arguably *are* such words. To use a well-worn example, the name 'Aristotle' is a type uses of which refer variously to the philosopher or to the shipping magnate. If proper names can be analysed as indexicals (and, in the footnote about Pelczar, Papineau acknowledges that they can), there is no reason why encyclopedia entries could not be.

In *Direct Reference,* I emphasized one potential consequence of the indexical view of names:

---

[3] If the presupposition fails, there may still be speaker's reference but there is no semantic reference.

'An ordinary indexical is associated with a particular relation R by the semantic conventions of the language. For every expression-type belonging to the category of ordinary indexicals, there is a specific relation R, such that the reference of a token of the expression is the object (or an object) which stands in that relation to the token. Different indexicals are thus paired with different relations by the semantic rules of the language. But it seems that all proper names are associated with the same relation R by the semantic rules of the language: the reference of a proper name, in all cases, is the entity which bears that name. In other words, while there is a distinct semantic rule for each indexical (the rule that 'I' refers to the speaker, 'you' to the addressee, 'this' to an entity contextually salient, and so forth), there is a single semantic rule for all proper names, namely the rule that a proper name refers to its bearer.' (Recanati 1993: 142)

Note that this is exactly like the case of encyclopedia entries as described by Papineau. Different encyclopedia entries are tokens of the same abstract type 'encyclopedia entry'. Similarly, in *Direct Reference*, I toyed with the idea that 'there is a single abstract indexical, call it PN, corresponding to all proper names, or to the general category of proper names. Every proper name would be an instance of this abstract indexical, which could thus be defined by the following rule: an instance of PN refers to the bearer of that instance' (Recanati 1993: 142).

Papineau also mentions a couple of cases in which, even though an indexical word is used in expressing a thought, there is no corresponding indexical constituent in the thought. He uses 'you' as an example. When I see John and have a thought about him that I want to communicate to him, I say 'you'. My thought — the thought I express — involves a lasting file about John. As Papineau says,

'I don't need to form an extra *you* file when I am about to address John, beyond any files about him I already have. So there is no reason to suppose that my utterance 'You gave a good talk yesterday' expresses some corresponding indexical *you* thought. Rather it is just the linguistic means that I use to express a pre-existing non-second-person thought.' (Papineau: 165)

Of course, the use of the linguistic means in question (the word 'you') to express a thought about John presupposes that John is identified as *the current addressee*. But according to Papineau,

'The mental type in question operates mainly in the sub-personal speech production system. Once this speech-production system 'knows' that *John is the current addressee*, it will set itself to express all

> John-referring thoughts using the word 'you'. In the normal case, I
> don't need to think about how to express my thoughts. My selection of
> words is generated automatically, courtesy of an automatic and uncon-
> scious system that figures out what grammatical string of words will
> best serve to express my thought in the current context.' (Papineau:
> 166)

This raises a very interesting issue — are there indexical files in
personal-level thought? — which I am about to consider. For the time
being, I simply note that Papineau's observation about 'you' is com-
patible with my account. The case discussed by Papineau is (in my
framework) a case in which the addressee is thought of under a com-
posite file resting on a number of ER relations to John, the person
the speaker is addressing. This is not a pure 'you' file. (Assuming
such files exist, they would be a variety of demonstrative files.)

I now turn to Papineau's second claim, and his most significant
objection to my account. Papineau thinks that as soon as we en-
counter an object we open a potentially lasting mental file about it,
abstracting away from our current ER relations to that object. He
does not think we need, in addition to such files (encyclopedia en-
tries), the fine-grained files based on particular ER relations, such
as demonstrative files. He claims there are no such files: the demon-
strative component in thought corresponds to egocentric features of
objects which are important because they guide action but which do
not serve to accumulate information about the objects. To accumu-
late information, we need a potentially lasting file.

For Papineau there is a subpersonal system of action-guidance
which, like the subpersonal system of speech production, needs in-
dexical categories; but *thought* is said to rest on a system of (endur-
ing) files which have the distinct function of accumulating informa-
tion about objects. It is a mistake to confuse the indexical categories
at work in the action-guidance system and the referential categories
(the files) we use in thought:

> '[The] automatic motor control system is not in the business of stor-
> ing information about the things it refers to, and so will not have any
> information-accumulating files associated with its tokens of *that thing
> there*.' (Papineau: 172)

Papineau's theory resembles a view put forward by Perry in vari-
ous places. Perry describes the mind as having both 'buffers' which

register egocentric information about the objects which play certain important epistemic-pragmatic roles in the subject's context (e.g. the role of current addressee), and 'detached files' about the objects in question. The buffers are arguably subpersonal, but this does not mean that we should only admit detached files as thought constituents, as Papineau suggests. In *Direct Reference*, after discussing Perry I offered a three-level picture with an intermediate category of *indexical object-files* between the buffers (bottom level) and the encyclopedia entries (top level).

One merit of having indexical files (in the narrow sense) as well as encyclopedia entries in thought is that this provides a bridge with the domain of perception, where reference goes through object files based on various tracking relations. The mechanism of incremental conversion makes it possible for indexical object-files to evolve into more and more abstract files, up to the encyclopedia entries which are 'detached' (in one sense at least). I think the transition between the two types of representation is an important issue, and that it must be explicitly represented in the theory. My system honours this requirement.

But Papineau is right to emphasize the dynamic unity of files across incremental conversion. In line with that emphasis, I myself introduced the (coarse-grained) 'piles' in addition to the (fine-grained) files in chapter VII of Mental Files. A possibly better way of accommodating Papineau's insights within my system would be to pair each file, qua body of information putatively about a single object, with a referential index. The index can play the role of address or label for the file: at any given time *t*, the index will give access to the file currently bearing the index. But the innovation I am contemplating to make the system more dynamic is this: *in incremental conversion the index will be transmitted from the source file to the successor file*. The index, thus construed, can do all the work done by the lasting file in Papineau's framework. But the fine-grained files do not disappear, in contrast to what happens in Papineau's framework. The reason why we need them, and cannot do simply with coarse-grained files à la Papineau, is that modes of presentation can be as fine-grained as one wishes.

Take Papineau's example: I have a thought about John which I want to communicate to him, so I say 'you'. As Papineau points out,

my thought involves my *John* file — a lasting, composite file rather than a short-lived, demonstrative file. But if, suddenly, I come to entertain a doubt about the identity of the man in front of me, I may think: 'Is *that man* really John? What about the moustache?' And I may ask him: 'Are you John?' in such cases I need to deploy fine-grained files based on certain ER-relations and not others. I need a demonstrative file in addition to my preexisting *John* file. (In general, I take it that one can modulate one's files more or less at will, by giving prominence to certain epistemically rewarding relations and bracketing others. Dennett's paper 'Where Am I?' provides nice examples of such modulation involving the 'self'-file, in contexts in which the presupposition of unity of various informational channels fails.)

To account for the interrogative thought 'Is that man really John?', Papineau will need to appeal to two potentially lasting files: my John file, and the new file created for the man I see, who may or may not be John. The new file is 'potentially lasting' (like all files in Papineau's framework), but it will actually last only if I keep using it; if the man I see turns out to be John I will discard the new file and not use it again, so it will decay and disappear. It will not last.

But I think it is not enough to posit two files to account for the case. We also need to account for the two *types* of mode of presentation at stake, corresponding to the singular terms 'that man' and 'John' respectively. The first mode of presentation has to do with our current (perceptual) relation to the man, while the other one is based on multiple information sources. I take the first mode of presentation, expressed by 'that man', to be a demonstrative file, and the other one, expressed by 'John', to be an encyclopedia entry.

As we have seen, Papineau rejects the very idea of a demonstrative file. He acknowledges the existence of indexical modes of presentation in thought (at the interface with the action-guiding system), but thinks such modes of presentation are not in the business of storing information about the referents: the accumulation of information task is incumbent upon the lasting, coarse-grained files. That, I think, is the weak point in Papineau's argument. In *Mental Files* I argued, contra Papineau, that demonstrative files can themselves be used to accumulate information during the thought episode which is their lifespan (and within the bounds of which they can be re-used).

Insofar as I can tell, there is no reason to restrict the function of storing information to lasting files.

4 Linking, merging, and coreference *de jure*

I take two terms to be coreferential *de jure* when they are associated with the same mental file. In identity judgements the terms flanking the identity sign are typically associated with distinct files. But we also need a notion of coreference *de jure* between pieces of information, to capture the idea that sometimes — when the subject 'trades upon identity' (Campbell) — it is presupposed that two pieces of information concern one and the same object. It would be natural to say that two pieces of information are coreferential *de jure* just in case they are 'co-filed' (to use Goodsell's terminology), i.e. belong to the same mental file. Yet that is not possible in my framework. Corresponding to identity judgments, there is the operation on files I call 'linking'. Linking two files makes the transfer of information possible between them. So, when you discover that A is B, you are licensed to transfer information in the A file into the B file, and vice versa. But that means that the files will now contain information gathered 'from outside', i.e., not through the ER-relation on which the file is based but through linking. I thus distinguish the 'nucleus' and the 'periphery' of the file. Only bits of information in the nucleus will be coreferential *de jure*. (Information in the periphery is only there because of a judgement of identity.) In other words: in my framework co-filing does not require presumptions of identity with presuppositional status. Defeasible identity judgments are sufficient to ground co-filing. It follows that co-filing is a weaker notion than coreference *de jure* between pieces of information: not all pieces of information that are co-filed are coreferential *de jure*.

Goodsell objects that the resulting theory is incoherent:

> 'Recanati explicitly claims that it is not the case that if *i* and *j* are in the same mental file, then *i* and *j* are de jure coreferential (94-95). This claim… seems inconsistent with files playing the mode of presentation role.' (Goodsell: 185-6)

Indeed modes of presentation determine reference, and they account for cognitive significance. Because modes of presentation play these two roles, Goodsell argues, it is a priori that two occurrences of the

same mode of presentation will corefer if they refer at all. So if two pieces of information are associated with the same mode of presentation (the same file), it will be presupposed that they concern the same object (if any). Thus it seems impossible to separate co-filing and coreference *de jure*.

I agree with Goodsell that there is a potential problem here. Fortunately, as she points out, another option is available. Moreover, it is an option which I think we need to consider if, following Pryor's interesting suggestions for developing the mental-file account, we opt for a graph-theoretic representation of relations (Pryor, forthcoming).

The idea I am now toying with is that linking should *not* be described as allowing 'information transfer' between files. Suppose we start with two files, A and B, and the identity of their referents is discovered (or is thought to have been discovered). Linking takes place, corresponding to the identity judgement 'A = B'. Linking enables information in one file to 'mix' with information in the other file in reasoning; but no transfer of information between files has to occur, strictly speaking. Linking does no more or no less than an identity judgment does.

Interestingly, Goodsell points out that

> 'It is possible that though *i* and *j* were initially treated as about the same thing in virtue of an identity judgement, that identity judgement becomes so embedded in the thinker's reasoning that she becomes disposed to reason as if *i* and *j* are about the same thing without deploying an additional identity premise in her reasoning.' (Goodsell: 187)

I think this is correct: an identity may acquire presuppositional status with time. The more we use the files A and B in tandem in reasoning (because of the identity belief which enables information in them to mix), the more we tend to presuppose the identity which grounds the coordinated use of the two files. At some point in the process identity will be presupposed. This point is often represented through the 'merging' idea: instead of two files, we only have one. A slightly different representation is made possible by the idea of referential index I introduced in my response to Papineau. When identity between A and B is presupposed, the two files come to share the same referential index. That means that the two files are now treated as a single, composite file (in virtue of the principle that two distinct files

cannot bear the same referential index at the same time).

Goodsell points out that, before the presupposition stage, when the two files are merely linked,

> 'I retain my ability to think independently about that-heared bird and that-seen bird, that is, it remains possible that [they are] different birds.' (Goodsell: 183)

This is right, but I would go further. Even when the identity is pre-supposed I may retain that ability. If I discover that the identity is mistaken and that there are actually two birds (one I see and one I hear), I can deconstruct the composite file and restore the initial indices. When, as in this case, the files correspond to distinct information channels (here, distinct modalities), un-merging the files is a relatively simple matter of index-splitting.

This relates to an issue I discussed in my response to Papineau. We can modulate our files in response to incoming information, e.g. we can stop using a composite file and start using two more specific files instead. As Goodsell points out in a footnote, 'to suppose that i and j are about different things, the thinker must move i and j to distinct files' (Goodsell: 185, note). The distribution of referential indices tracks the current presuppositional state of the thinker, but that state dynamically evolves. Presuppositions come in and go out of existence all the time.

5 Mental files vs. competing accounts

In *Mental Files* I discuss alternative frameworks such as two-dimensional neo-Descriptivism, token-reflexivism, and Lewis's centered-content framework. Lawlor argues that I am too critical of these frameworks, and that both the token-reflexive framework and the descriptivist framework have the resources to address my worries in a potentially satisfactory manner. In a syncretic spirit, she claims that 'one does not need to argue for the absolute untenability of other approaches in order to advocate in favor of the files framework' (Lawlor: 156). Likewise, García-Carpintero attempts to defend a view similar to Perry's (involving a distinction between the singular content of a state and its primary content, understood as a token-reflexive proposition), and suggests that it is compatible with my ap-

proach in terms of files, though not with the Lewisian approach I advocate in *Perspectival Thought* and my papers on IEM. (On the last point, see the next section.)

But I do not think I ever argued for the absolute untenability of the other approaches. On the contrary, I hold that all the main approaches shed some light on the phenomenon, and that each improves as a result of elaboration made necessary by criticism emanating from the other approaches (including, hopefully, criticism presented in *Mental Files*). I believe that, in the end, there is a true convergence, which I tried to highlight in the last chapter of the book.

Lawlor speaks as if my mental files account stood in contrast to Perry's account, but it doesn't; it's an elaboration of it. Perry distinguishes between the content (a singular proposition) and the belief state, and assigns the state a primary content akin to a kaplanian character. Belief states are vehicles, in my terminology, and Perry himself appeals to files qua mental particulars in analysing them. My theory has exactly the same ingredients as his. Note that, in Perry's work, you find elements from different frameworks: a two-dimensional component, a mental-particular component, a token-reflexive component (not to mention the Lewisian component which surfaces in 'Thought without Representation').

Even though García-Carpintero thinks my mental files account is in tension with the Lewisian approach of *Perspectival Thought*, he acknowledges that elements from different frameworks can be combined. Thus he takes Stalnakerian diagonal propositions, or token-reflexive propositions of the sort he himself advocates, to be less an alternative to Perry's initial theory of indexical belief than an elaboration:

> 'Perry accepts that, for the kind of consideration about informational content that Stalnaker pointed out, (…) token-reflexive contents provide a better representation of the significance of belief-states than the one he had earlier suggested in terms of Kaplanian characters. (…) On this interpretation, the proposal is just a refined way of understanding the significance of belief-states; but an adequate account of *de se* contents (hence of the nature of attitudes and speech acts in general) still requires the distinction between belief-contents and belief-states (ways of accessing the content). The modification of Perry's original proposal lies only in that now the significance of belief-states is characterized in the traditional propositional way that token-reflexive contents afford.' (García-Carpintero: 196-7)

Similarly, Stalnaker points out, Lewis's centred contents can be used to model the primary content of belief states in Perry's framework (Stalnaker 2003: note 255).

Let me say a bit more about the convergence I talked about. The most important thing is that everybody agrees that we need two levels of content. The distinction can be captured in different ways (character/content; diagonal/horizontal; reflexive content/subject matter content, etc.), but it is unescapable. The only (apparent) exception is Lewis: he seems to have a single level of content — centred contents. But this is an illusion (Recanati 2012b: 249). Lewis's centred contents are relativized propositions, and relativized propositions can only be evaluated against an appropriate index. If you pair the index and the centred content, you get an Austinian proposition with classical truth-conditions. Suppose the indexical content of a belief state is a property *P* which the believer self-ascribes. Then the belief is true iff *the believer has P*. The right-hand side of the biconditional gives us the secondary content of the belief, while the property *P* itself corresponds to its primary content. So, as I emphasized in *Perspectival Thought*, there *are* two levels of content in the centred content framework.

The main problem with Relational Descriptivism (i.e. the version of Descriptivism which takes on board causal/relational factors and rephrases the descriptive contents in terms of them) is that it puts the acquaintance relations into the content of the belief even though they clearly don't belong to its subject matter. But if we add to Relational Descriptivism a distinction between two levels of content — as 2-D Relational Descriptism does — the objection no longer arises. The acquaintance relations are now represented as part of *primary* content, not as part of secondary or subject-matter content.

In *Mental Files*, I objected to all 'internalization' of acquaintance relations, on grounds of intellectualism. Acquaintance relations are determinative of content, but they are not themselves represented. This must be qualified, to take account of the distinction between primary and secondary content. Acquaintance relations are not (explicitly) represented as part of secondary content, but that is compatible with their being, perhaps, *implicitly* represented as part of primary content. As Lawlor points out, 'unreflective people reason in ways that are sensitive to the mode by which they acquire in-

formation' (153); that sensitivity is not the same thing as having 'a higher-order thought about the sources of one's information' (*id*.). I agree. I also agree that distinguishing between unreflective sensitivity to ER-relations and explicit representation of them makes it possible for 2-D Relational Descriptivism to evade the charge of intellectualism, while still somehow incorporating the ER relations into primary content. But that means that, just as there are two levels of content in that framework, there are two distinct 'grasping' relations corresponding to them. The secondary content is what the thinker (explicitly) represents, but the primary content is not represented in the same way. The thinker stands in a different relation to primary content than the relation he or she stands in to secondary content. Modulo this distinction between two grasping relations, 2-D Relational Descriptivism can be saved.

But if that is so, then, Lawlor argues, 'it seems we do not need a file framework to dispense with the relevant Descriptivist commitment. So what recommends a files-based approach?' (154) As I pointed out already, the mental-file account I offer seeks to integrate as much as possible of the insights underlying the other accounts; but, of course, it adds something to them. Why add something? Because it's not enough to posit two levels of content, even if you add a corresponding distinction between two grasping relations, one appropriate to secondary content and the other one to primary content. You must say something about *what the relation is* in the case of primary content.

Perry, one of the few authors who've touched upon this issue, calls that relation 'attunement'. He describes it as follows:

> 'Attunement to the relation that our self-notions have to ourselves, or our perceptions have to the object they are of, does not require belief or thought about the relation; it requires know-how, not knowledge that.' (Perry 2012: 99)

This is very much the Lawlor point about unreflective sensitivity vs explicit representation, but Perry more specifically talks of 'know how'. What that means, according to the gloss I offer in *Mental Files*, is that primary contents (which Perry initially modeled as Kaplanian characters)

> 'have a *procedural* nature. They correspond to certain functions which words or mental vehicles have. The functions in question are not rep-

resented. The vehicles simply *have* those functions and they operate in context according to these functions. The referential content of the vehicle depends upon that operation' (Recanati 2012b: 248)

On this view, primary content must be cashed out in terms of functions or roles. The functions are not represented or thought about by the users, yet (as Lawlor points out) they are transparent to them in the sense that they 'know how' to operate according to them. Now, there is no function or role unless it is the function/role of *something*. That something is the vehicle. The mental file account completes the story by adding the vehicles and interpreting primary content in terms of their functional roles.

In the last chapter of *Mental Files* I offer a critical assessment of token-reflexivism, which Lawlor finds too harsh. But she misinterprets me. My goal is not to argue against token-reflexivism, but to show where it leads. I argue that, if you embrace token-reflexivism and try to meet the objections it raises as they come up, you end up with something like the theory of mental files. The argument proceeds in three steps. I show first that token-reflexivism, properly understood, reduces to a view very similar to Lewis's. Then I show that this view (Lewisian token-reflexivism, as we may call it) itself raises objections — the same which Lewis's original framework raises — but that these objections can be met by shifting to a multi-centred variant. Finally I show that (suitably elaborated in order to meet a new set of objections) the multi-centred variant leads us to the theory of mental files.

As we have seen already, García-Carpintero holds that token-reflexive propositions can usefully replace the kaplanian characters which Perry initially appealed to to model the primary content of doxastic vehicles.[4] García-Carpintero defends a version of the Perry view on which the doxastic vehicle has both a referential content (a singular proposition) and a reflexive content which captures the significance of the state (its primary content). In this framework, the mode of presentation under which the object a singular belief is about is thought of can be rendered by a token-reflexive description.

---

[4] García-Carpintero, following Stalnaker, argues that token-reflexive propositions are better suited than characters or centered contents to explain communication.

So if I think 'I am tired', the primary content of my thought occurrence (call it *u*) is the token-reflexive proposition that the thinker of *u* is tired at the time of *u* (in the world of *u*). Similar views have been argued for by John Searle, James Higginbotham, and John Perry.

This view raises a prima facie problem. Under which mode of presentation is the occurrence itself, *u*, thought about when we think the thought? The mode of presentation of *u* can't be a token-reflexive description, on pains of circularity. The solution consists in arguing that *u* itself is not 'thought about' in the way the objects in the subject-matter content are thought about. Ordinary objects are thought of under modes of presentation which (according to the view) can be cashed out in token-reflexive terms; but the mental occurrences which feature in the token-reflexive descriptions ('thinker of *u*', 'time of *u*', etc.) are not themselves thought of under modes of presentation. They are directly given, in the flesh. This is the old Russellian idea of (strong) acquaintance or 'super-direct reference' (as I call it in the book), an idea which Perry endorses. Now this idea can be elegantly expressed in a variant of Lewis's framework, by centering the content on the mental occurrence of which it is the content:

> 'This idea can be couched in Lewis's framework, by externalizing the occurrence *u* and letting it be directly provided by the context. Everything is then described relative to *u*, but *u* itself is *given*, it is not represented. On this mixture of the two frameworks (centered worlds and reflexivism), the content of a mental occurrence is a *property of occurrences*, and that content is evaluated with respect to a contextual index containing the occurrence itself. On this Lewis-inspired view, to judge something by assertively tokening a certain representation is to ascribe to the token the property that is its content. Here reflexivity is guaranteed by the pragmatic architecture of the act of judgment. So when you think 'I am tired', the content of the thought is the property an occurrence has just in case the thinker of that occurrence is tired at the time of the occurrence in the world of the occurrence. To think the thought (or to think it assertively) is to ascribe that property to the current occurrence *u* you are producing.' (Recanati 2012b: 253)

However, both Lewis's original theory and the token-reflexive variant I have just presented raise the 'Cartesian Asymmetry' objection:

> 'the main problem (…) is the asymmetry between different objects of thought. Everything is thought of descriptively, except for a single element which is externalized and serves as *universal anchor* for all the content. Although I have no knock down argument against this approach, I find it unsufficiently motivated and too much in the grip of a rather

extreme Cartesian picture. Why not appeal instead to *multiple* anchors, corresponding to all the acquaintance relations in which we stand to objects of thought?' (Recanati 2012b: 253)

On the mental file picture, there is no asymmetry. Singular thoughts are about objects that are all represented under nondescriptive modes of presentation based on acquaintance relations. There is no privileged subset of objects (the subject-at-a-time, or his/her mental occurrences) which can be thought of 'super-directly' and in terms of which all the other objects are described.

Among the Lewisians, some have tried to get rid of the Cartesian asymmetry by substituting multi-centred worlds for Lewis's centred worlds. *All* the objects of thought can be externalized and fed into the index of evaluation, in a multi-centred framework. I have much sympathy for this framework, but in the book I introduce a friendly amendment: instead of including a sequence of objects (Ninan's '*res*-sequence') in the 'base world' serving as index of evaluation, I propose to include a sequence of files. This makes it possible to account for Frege cases and empty cases.[5] (The objects themselves can be easily retrieved: the objects of thought are the referents of the files in the base world, i.e. those objects which, as a matter of contingent fact, stand in the right relations to the files.) This is a 'Lewisian' version of the mental file framework. The main insight behind token-reflexivism is not lost, however. What fixes the reference of our files are certain relations between the files (tokens) and objects in the external environment, relations to which thinkers are 'attuned' and which are presupposed when the files are deployed.

Coliva and Belleri worry that my friendly amendment to the multi-centred world framework may be too drastic:

'Sequences of individuals are expunged from centred worlds, and only mental files are kept (256, 258). If this is so, then the files seem to acquire a strange status. On the one hand, they are mental, 'internal' objects, which act as vehicles of thought or 'mental singular terms' (viii, 35, 182, 244-5); on the other hand, they are the 'anchors' of our *de re* thoughts (253). These two features, however, seem difficult to reconcile: for one would think that the objects of our *de re* thoughts are external to the mind, and that they do not coincide with the vehicles we

[5] Ninan himself introduces acquaintance relations, in additions to the objects in the *res*-sequence, to account for Frege cases. But to deal with the empty case and other cases in which there is no acquaintance, you need the files.

use to refer to things in the world. Recanati could reply that the proper objects of our *de re* thoughts are really the referents of the mental files at issue — not the files themselves. This, however, just suggests that mental files alone are not enough in order to capture *de re* thought: individuals matter as well, and they should find their place in a suitable semantics for this kind of phenomenon.' (Coliva and Belleri: 115-6)

Of course, individuals matter, since individuals are what singular thoughts are about. But real individuals only feature in the thought's *secondary* content. The thought's secondary content depends upon the environment. In the world in which the subject thinks the thought (the base world) the subject's files are suitably related to individuals who, in virtue of these relations, count as the referents of the files. These referents feature in the thought's secondary content, and have a place in the semantics. So singularity and world-involvingness are not lost, in that framework. But if we are concerned with the thought's *primary* content, we can capture it without bringing real individuals into the picture. The thought's primary content is a *property* of a sequence of objects, representable as a set of multi-centred worlds. These worlds are the subject's doxastic alternatives. In each of the subject's doxastic alternatives, there will be a sequence of objects (possibly different objects in different worlds) corresponding to the mental files tokened in the base-world. Which objects the subject's thought actually is about will be determined by objective features of the base-world: they will be the objects which the files refer to in that world (in virtue of the ER-relations holding at that world). So: mental files are, indeed, not sufficient to capture (the secondary content of) *de re* thought: the world has to play its part. But they are sufficient to capture the thought's primary content.

6 IEM and the *de se*

García-Carpintero detects a tension between my 'Lewisian' account of *de se* content in *Perspectival Thought* and the theory of mental files. He takes the latter to be compatible with his own token-reflexive approach, which uses only classical propositions (both at the primary and the secondary level). I agree that they are compatible, but I don't think the token-reflexive view has to be seen as competing with the centred-content approach. As I pointed out in section 5, the token-

reflexive view *can* be formulated in Lewis's framework, by centering the contents on the occurrences of which they are the content.

In *Perspectival Thought* I presented my own take on the primary/ secondary distinction. I distinguished two levels. The first one is that of 'explicit' content. I call that primary content the *lekton*. The *lekton* is a centred content (a 'relativized proposition'): it can only be evaluated as true or false against an appropriate circumstance of evaluation involving not only a possible world, but also an additional component corresponding to that on which the content is 'centred': the subject at a time in Lewis's original framework, the occurrence in Lewisian token-reflexivism, the topic situation in situation semantics, or whatever. That component is not explicitly represented in the *lekton* because it is contextually given or taken for granted. Thus the perception-based thought 'it is raining' has a content that is a property of situations (the property a situation has just in case it is raining in that situation). The situation the thought is about is not explicitly represented (as it would be if the thought were 'it is raining *here*'), but it is determined as the relevant situation of evaluation by the perceptual nature of the experience which grounds the thought: perception is perception of the local situation, so the judgment 'it is raining', based on perception, is bound to concern the local situation (even though the latter is not explicitly represented in the content). The judgment is true iff *it is raining in that local situation*. The truth-conditions in the right-hand-side of the biconditional correspond to the secondary content of the thought. The secondary content is classical (uncentered) and results from the interaction of the explicit content, which is centred, and the relevant circumstance of evaluation. (This interaction is what the notion of 'Austinian proposition' is meant to capture.)

The central insight here comes from Perry's 'Thought without Representation', where he argues that sometimes, 'architectural or external constraints make internal representation unnecessary' (Perry 1993: 221). I use that feature to account for immunity to error through misidentification. When you detect the position of your legs through proprioception, the fact that it is *your* legs whose position is represented in the content of the proprioceptive experience transparently follows from the fact that the experience is proprioceptive. Proprioceptive experience is bound to concern the experi-

encer's own body. Still, the content of the experience does not have to represent the subject whose body is in question; it does not have to *identify* the subject whose bodily condition is represented. It is the (proprioceptive) *mode* of the experience which determines that its content (a bodily property) concerns the subject who is having the experience, and not some other subject. There is, as Perry puts it, an 'architectural constraint' that the body which a proprioceptive experience concerns is the subject's own body. In virtue of that constraint, the issue of whose legs are crossed simply does not arise. Misidentification is impossible because there is no identification in the first place.

To be sure, the proprioceptive experience is veridical if and only if *the subject's* legs are crossed. This corresponds to the secondary content of the state — its truth-conditions. But the subject is an 'unarticulated' constituent of that secondary content. It is unarticulated because it is not explicitly represented in the primary content of the state, but contributed by the proprioceptive mode (through the circumstancial component it determines).

García-Carpintero contrasts this account with the so-called 'Simple Account', according to which immunity to error through misidentification is a negative property: there is IEM whenever the first person judgment 'I am *F*' does not include, among its epistemic grounds, an identity premise 'I = *a*' (where '*a*' refers to an individual independently thought to be *F*). He says that I offer my account 'instead of' the Simple Account,[6] and adjudicates in favor of the latter. But these accounts do not compete. My account explains *why*, in proprioceptive cases and other cases like it, there is no need for an identity premise in the grounds: it is the mode of the experience which determines what it's about (or more precisely: what it 'concerns').[7]

---

[6] 'The explanation that the Simple Account affords appeals to the absence of an identity claim in the justificational structure; Recanati's explanation appeals instead to the absence of a conception of the self in the content of the IEM judgments' (García-Carpintero: 201).

[7] See Wright (2012: 273): 'Recanati's core proposal is, in a way, perfectly consistent with the Simple Account. (…) [It] can be viewed as… an attempt to characterize the distinctive justificational architecture of those I-thoughts that are IEM in a fashion that does indeed explain why the Simple Account applies'.

Following other authors, García-Carpintero distinguishes two senses of 'identification': 'identification$_C$' means that the subject is explicitly represented in the content of the judgment; 'identification$_P$' means that there is an identity premise 'I = a' in the judgement's grounds. He rightly points out that there can be identification$_C$ without identification$_P$, and that in such cases the judgment is still IEM. (I will discuss these cases in a minute.) But I don't think this supports the Simple Account as opposed to mine. What my account says is simply this: sometimes, there is no identification$_C$ of the subject in the *lekton* because an 'architectural' mechanism ensures that the content of the state is about the subject's own body. Using Searle's mode/content distinction: in the relevant cases it is the mode of the experience, not its content, which is responsible for its first personal character, by fixing the self as the relevant point of evaluation for the content. Since, in such cases, the mode dictates that the content can only be self-ascribed, no identity premise (no identification$_P$) is needed to ground the first person judgment. There is no sense in which I can think 'Someone is *F*, but is it me?', because the mode of the experience which grounds the judgment precludes the possibility that it might not be me. This is entirely compatible with the Simple Account. Again, I explain *why*, in the relevant cases, no identity premise is involved.

García-Carpintero cites two objections which have been made to my account. The first is that there can be IEM even if the subject is explicitly represented in the content of the judgment. This corresponds to the cases in which there is identification$_C$ without identification$_P$. As García-Carpintero points out, I acknowledge their existence, and I account for them by appealing to the process of 'reflection' through which an element of secondary content implicitly contributed by the mode is made explicit and gets represented in the primary content of the reflective judgment (Recanati 2012a). Reflection is, for example, the transition from 'It is raining' to 'It is raining here'. This transition never involves adding anything to the grounds of the original judgment. If I am justified in believing, on the basis of my perception, 'it is raining', then I am justified in believing, on the same basis, 'it is raining here'. Because the grounds of the post-reflection judgment are the same as the grounds of the original judgment, if the latter is IEM by the simple account, i.e. lacks an identity premise

among its grounds, then the former is bound to be IEM too. So I have no trouble with cases of identification$_C$ without identification$_P$.

The other objection is this. Immunity to error through misidentification also affects judgments that are not in the first person. A demonstrative judgment 'That is *F*' can be (and typically is) IEM, by the lights of most authors who write about the topic. But my account, according to García-Carpintero (following Wright), fails to extend to demonstrative IEM. Now, I don't agree that it fails, though I agree that some work has to be done to implement my proposal in the demonstrative case. But even if my account failed to extend to demonstrative IEM, that would not necessarily be a problem since the Simple Account takes IEM to be a negative property. As Wright himself suggested, different mechanisms can be at work in the first person case and in the demonstrative case; the explanation of why an identity premise is not required may be different in the two cases (Wright 2012: 274).

I now turn to the tension which García-Carpintero detects between my Lewisian account of IEM and the mental file framework. Mental files are singular terms in the language of thought, so if first person thoughts are thoughts involving the Self file, then their content is not Lewisian: it is not selfless. I quote García-Carpintero:

> 'Simply put, mental files are individual concepts, and concepts are ingredients of contents. If the account of *de se* thoughts is that they deploy the Self mental file, then the contents of *de se* thoughts have this concept as an ingredient. (…) In the terminology used above, Recanati's account of *de se* thoughts and their communication appears to involve an identification$_C$ — the Self file.' (Garcia-Carpintero: 203)

This is a good objection, and I am happy of the opportunity it provides to clarify my views on this topic (the relations between the mental file account and the Lewisian account in *Perspectival Thought*).

García-Carpintero notes that, in addition to files (which are conceptual and satisfy the Generality Constraint), I also posit 'proto-files' which are nonconceptual. Thus there is a proto-file SELF*, which can *only* host information gained in the first person way (while the SELF file can host any information one gains about oneself). But this does not neatly correlate with the distinction between 'explicit' *de se* thoughts, in which the self is represented, and 'implicit' *de se* thoughts whose content is selfless. It's OK to say that explicit *de se*

thoughts involve the SELF file, but what about implicit *de se* thoughts? Carpintero hypothesises that 'it is the SELF* proto-file that figures in basic, "implicit" *de se* contents', but objects that 'contentful states involving the SELF* proto-file still appear to involve an identification$_C$' (p. 10). Indeed, if the SELF* proto-file contributes to the primary content of first person experience at the nonconceptual level, then the primary content in question is *not* selfless. Hence García-Carpintero's doubts regarding 'the compatibility of [my] Lewisian account of de se contents and the mental files approach' (192).

Yet, I maintain, they *are* compatible. My Lewisian account is an account of the (selfless) content of first person *experience*. Based on the experience, there is a first person *thought* involving the SELF file. That thought is explicitly first personal. It involves identification$_C$ but no identification$_P$ because, through 'reflection', the thought inherits the grounds of the underlying experience.

What about the proto-file? In virtue of its mode, the experience feeds its selfless content into the SELF* proto-file, but I do not regard the proto-file as an aspect of the experience itself. The experience I analyse into (selfless) content and mode. Nor is the self* proto-file used in thought; rather, it is used in the guidance of action. So I maintain that the Lewisian account applies to first person experience, while the mental file account applies to reflective thought in which the self is explicitly represented. The transition from first person experience to explicit first person thought is only possible if the subject possesses a first-person concept, i.e. a SELF file.

In closing, let me address another worry, voiced by Coliva and Belleri, regarding my account of IEM:

> 'The identity between oneself and the person whose body is responsible for the proprioceptive/self-locating information one is receiving, or between oneself and the person whose memories one is storing, is only contingent. But if our SELF file should guarantee knowledge of its referent in all possible circumstances, for otherwise it would no longer be a SELF file, it cannot be based on those epistemically rewarding relations. For, in some circumstances, they would not deliver information about oneself… Hence, we need a relation which secures the knowledgeable identity of the subject to himself *in all possible circumstances*. That relation, we take it, would rather be the one between the subject and the thinker of a given occurrent thought. What this shows is that not all singular modes of presentation of a given entity, in this case the subject himself, are on a par with respect to a given file. Some

would seem to be constitutive of it, like 'the thinker of this occurrent thought' for SELF.' (Coliva and Belleri: 113-4)

But the reference of a file is not the object one is gaining information from through the relevant information channel. In the deviant cases Coliva and Belleri allude to (quasi-memory, quasi-perception, quasi-proprioception, etc.), the information channel delivers information whose (deviant) source turns out to be another individual. But *the file still refers to the subject under those circumstances*: the subject still refers to himself or herself by deploying the SELF file. The file refers to the object one stands in the right ER-relation to, and in the case of the SELF file, that relation is identity.

The information channels correspond to ways of gaining information about an object *x* that are normally available to the subject when, and only when, s/he stands in the right relation to *x*. The relation is said to be 'epistemically rewarding' because, when that relation holds, one is in a position to gain information about *x* in a special way, e.g. (in the first person case) 'from inside'.[8] That special information channel normally delivers information about the object *x* one stands in the corresponding ER relation to, but in the deviant cases it does not: in quasi-proprioception, for example, information gained from inside (through proprioception) has its source in another individual than the individual we are. Still, the SELF file refers to oneself, even in these circumstances, because it refers to the object one stands in the right relation to — not to the object we are gaining information from.

Be that as it may, Coliva's and Belleri's search for a relation that guarantees self-knowledge in all possible circumstances sounds suspiciously Cartesian and seems to fall prey to Wright's objection:

> 'Those of my thoughts that are immune to error through misidentification are not so because they involve super-sure identification of myself, conceived on the model of knowledge of an identity 'I am *a*' [e.g : I am

---

[8] The ER relation, in the first person case, is *not* the relation of 'gaining information about *x* from inside'. The ER relation is identity. The ER relation is one thing, and the information channel corresponding to it is another. Normally, they go together: one can gain information about *x* from inside just in case one is *x*. Yet the 'internal' information channel can fail to deliver information about oneself, while the SELF file, based upon the ER relation (identity), can't fail to refer to oneself.

the thinker of this thought — FR], but because no such judgment…
features in their justificational architecture.' (Wright 2012: 253)

This — the Simple Account — applies to all IEM thoughts. Now
there is an important distinction (emphasized by Coliva in her work)
between logical IEM and contingent or *de facto* IEM. The mechanism
I posit to explain IEM is meant to account for contingent first person
IEM. I have tried to extend it to another kind of contingent IEM,
namely demonstrative IEM — without being able to convince either
Wright or García-Carpintero that this is the right way to go. I suspect
that the underlying mechanism is different in the case of logical IEM,
a property exhibited by self-ascription of one's occurrent thoughts.
Given that (i) the SELF file refers to the subject in whose thought the
file is deployed, and (ii) the subject is 'attuned' to that fact, it would
be incoherent for a reflective subject to deny ownership of his or her
own occurrent thoughts. In this case, perhaps, García-Carpintero's
analysis of IEM applies.[9]

François Recanati
Institut Jean-Nicod
Ecole Normale Superieure
29 rue d'Ulm, 75005 Paris
France
recanati@ens.fr

References

Campbell, John. 1987. Is Sense Transparent? *Proceedings of the Aristotelian Society* 88: 273-92.

Coliva, Annalisa. 2006. Error Through Misidentification: Some Varieties. *Journal of Philosophy* 103: 403-25.

Dennett, Daniel. 1978. Where Am I? In *Brainstorms: Philosophical Essays on Mind and Psychology*, Montgomery, VT: Bradford Books and Hassocks, Sussex: Harvester.

Evans, Gareth. 1982. *The Varieties of Reference*. Oxford: Clarendon Press.

Fiengo, Robert and May, Robert. 2006. De Lingua *Belief*. Cambridge, Mass.: Bradford Books/MIT Press.

Grice, H. Paul. 1989. *Studies in the Way of Words*. Cambridge, Mass.: Harvard

---

[9] 'A thought of the form *A is P* is IEM when the ascribed property *P* [e.g. being
the thinker of this thought] is already ascribed to the referent of the concept *A* as
part of its background reference-fixing information' (García-Carpintero: 203).
In the case of the the SELF file, the reference-fixing rule is that it refers to the
thinker in whose thought the file is deployed.

University Press.

Kaplan, David. 1990. Words. *Proceedings of the Aristotelian Society*, Supp. Vol. 64: 93-119.

Kaplan, David. 1989a. Demonstratives. In *Themes from Kaplan*, ed. by Joseph Almog, Howard Wettstein and John Perry, 481-563. New York: Oxford University Press.

Kaplan, David. 1989b. Afterthoughts. In *Themes from Kaplan*, ed. by Joseph Almog, Howard Wettstein and John Perry, 565-616. New York: Oxford University Press.

Kripke, Saul A. 2013. *Reference and Existence*. New York: Oxford University Press.

Lewis, David. 1978. Truth in Fiction. *American Philosophical Quarterly* 15: 37-46.

Lewis, David. 1979. Attitudes De Dicto and De Se. *Philosophical Review* 88: 513-43.

Ninan, Dilip. 2008. *Imagination, Content, and the Self*. Ph. D. dissertation, MIT.

Papineau, David. 1986. Phenomenal and Perceptual Concepts. In *Phenomenal Concepts and Phenomenal Knowledge*, ed. by Torin Alter and Sven Walter, 111-44. New York: Oxford University Press.

Perry, John. 1993. *The Problem of the Essential Indexical, and Other Essays*. Oxford: Oxford University Press.

Perry, John. 2001. *Reference and Reflexivity*. Stanford : CSLI.

Perry, John. 2012. Thinking about the Self. In *Consciousness and the Self : New Essays*, ed. by JeeLoo Liu and John Perry, 76-100. Cambridge : Cambridge University Press.

Pryor, James. (Forthcoming.) Acquaintance, Mental Files and Mental Graphs.

Recanati, François. 1993. *Direct Reference : From Language to Thought*. Oxford: Blackwell.

Recanati, François. 2007. *Perspectival Thought*. Oxford: Clarendon Press.

Recanati, François. 2012a. Immunity to Error through Misidentification: What it is and Where it Comes From. In *Immunity to Error Through Misidentification: New Essays*, ed. by Simon Prosser and François Recanati, 180-201. Cambridge: Cambridge University Press.

Recanati, François. 2012b. *Mental Files*. Oxford: Oxford University Press.

Russell, Bertrand. 2010-11. Knowledge by Acquaintance and Knowledge by Description. *Proceedings of the Aristotelian Society* 11: 108-28.

Schiffer, Stephen. 1978. The Basis of Reference. *Erkenntnis* 13: 171-206.

Stalnaker, Robert. 2003. *Ways a World Might Be*. Oxford: Clarendon Press.

Walton, Kendal. 1990. *Mimesis as Make-Believe*. Cambridge: Harvard University Press.

Wright, Crispin. 2012. Reflections on François Recanati's 'Immunity to Error through Misidentification: What it Is and Where it Comes From'. In *Immunity to Error Through Misidentification: New Essays*, ed. by Simon Prosser and François Recanati, 247-80. Cambridge: Cambridge University Press.