

# Indexicality, Transparency, and Mental Files

Derek Ball  
University of St Andrews

It would be easy to read Francois Recanati's *Mental Files* as a sustained defence of the utility of a certain metaphor: that the mind is like a filing cabinet, and that our mental representations are like file cards on which our beliefs are written. But too much focus on the metaphor sells Recanati's project short; Recanati advances a number of challenging claims about mental representation, which, though phrased in terms of the file metaphor, are ultimately independent of it. This paper aims to challenge two of these claims: that mental representations are *indexical*, and that they are *transparent*. I begin by explaining the role indexicality plays in Recanati's system.

## 1 Mental Files and Indexicality

According to Recanati, beliefs are structured mental representations, and some of their components are usefully thought of as being like files that contain information. Recanati claims that each file is associated with an acquaintance relation (which Recanati calls an epistemically rewarding (ER) relation). The referent of the file is the object that stands in the relation to the file. So, for example, some files – the *HERE* files – are associated with a “relation which holds between a place and a mental file whenever the mental file is tokened in the mind of a person who occupies the place and serves to store information that person is in a position to gain in virtue of occupying that place” (2012, p. 70); others – the *SELF* files – are associated with a relation that holds between a person and a file just in case the file is tokened in the person's mind and serves to store information that the person is in a position to gain about herself in virtue of the fact that she is herself (2012, p. 70). Referring expressions in language inherit their reference from the mental files with which they are associated (2012, p. xiii).

Given this view of the reference-fixing role of ER-relations, we can state token-reflexive rules by which the referent of a file can be determined (2012, pp. 69-70). For example, the referent of a *HERE* file is the place that the person who has tokened that file is in. Recanati regards these token-reflexive relations as capturing what all files of a given type have in common. This is the sense in which files are *indexical* (2012, pp. 59-60).

On Recanati's view, *all* files are indexical in this way – even those (like *SELF* files, or *stable files* (the “encyclopedia entries” that correspond to names)) that do not ordinarily shift their referents. There is a terminological issue here about whether this is a useful way to use the word “indexical”, but let's grant Recanati his usage. It still seems clear that there is a theoretically interesting difference between (say) names on the one hand, and words like “you” and “that” (and, on Recanati's view, their associated files) on the other. As a rough first shot, the token-reflexive rule for “you” and “that” will link tokens of these terms will very often link different tokens to different referents, even if we hold the speaker fixed. Call terms and mental files that shift reference in this way *practically indexical*.

On Recanati's view, many mental files are not practically indexical. Notable among these are the stable files. Stable files are linked to a higher-order ER relation: the relation that holds between an object and a file just in case some other ER relation holds between the object and the file. But, Recanati claims, these rest on a base of files (or proto-files, which are based on a particular ER relation, but (unlike full-fledged files) can only contain information obtained through that relation) that are closely linked to specific ER relations. Recanati writes:

files at each of the levels I have described presuppose files at the previous level. Proto-files are the most basic; conceptual files are generated from them (through what I called ‘expansion’). Among conceptual files, first-order files are more basic, since higher-order files—encyclopedia entries—presuppose them. (2012, p. 75)

Since – with a few exceptions, such as the relation associated with the SELF file – ER relations come and go, many or most of these more basic files will be practically indexical. So Recanati endorses the following claim:

**The Hierarchy Thesis** Stable “encyclopedia entry” files can exist only if practically indexical files exist.

I said earlier that the referent of a mental file is the object that stands to the file in the relation associated with the file. What happens if more than one object stands in the relation to the file (or if different objects stand in the relation to the file at different times?) On Recanati’s view, files are subject to a norm of proper functioning:

**The ER Norm** A file should be maintained only if exactly one object stands to it in its associated ER relation, and only for as long as the same object stands in this relation.

In general, if this norm is not satisfied, the file fails to refer. So, for example, if one has a HERE file, and one’s location changes, in order to satisfy the norm one must remove the file (presumably copying the information contained in it to a file of another type), or change the file into a file of another type. One can then form another HERE file, which will refer to one’s new location.

The situation is somewhat more complicated with respect to stable files, since these are associated with multiple ER relations. In the case where such a file comes to bear different ER relations to different individuals, Recanati claims that the referent is the *dominant* source of information in the file (2012, p. 140). Recanati leaves the notion of dominance largely intuitive, but he does hold that dominance may shift depending on what information in the file is most “strongly activated” by the purposes at hand (2012, p. 140). So in this sort of case, the referent of a stable file exhibits a sort of context-sensitivity: it will refer to different individuals depending on what part of its information is most relevant in a given situation. And if no single source of information is dominant in a given situation, then the file fails to refer in that situation.

A system of mental files must be designed in such a way that it can obey the ER Norm, at least in typical cases. I see two ways such a design could go. Let me introduce them by discussing two ways one might design a computer to store information about different times. Computer clocks depend on a timer chip which sends a signal at a regular frequency (for example, on typical PCs, 18.2 times per second). In typical computers, this signal advances a counter. This counter represents the present time.

Now suppose that the computer with this sort of clock needed to generate files associated with different times. On Recanati's picture, the system should work like this: in addition to a series of stable files – the encyclopedia entries about times, or the names of times – it would have a buffer file in which it collects information about the present moment. At each signal from the timer chip, it would generate a new, stable file, copy the information from the buffer into the stable file, and remove the information from the buffer. The buffer file is naturally regarded as a sort of indexical – a NOW file. Call a computer that works in this way a *practical-indexicals-basic* – or for short *indexicals<sub>pr</sub>-basic* – system.

Surely there could be a system of this kind. But there is another way that a computer could keep track of information about various times (without any kind of NOW buffer). Recall that computers increment a counter at each signal from the timer chip. A computer could be designed so that just after each signal, it creates a new file, labeled with the value from the counter. Information about the present time could be stored in the highest-valued file; information about past times could be stored in relevant lower valued files.

It is most natural to regard the files used by such a system as stable, encyclopedia entry-style files rather than as practically indexical files. A file might be used just after it is created to store information about the present time; but the file is retained even after a new, higher-valued file is created, and even from the beginning it could have as the function of storing information from any number of ER relations. So, at least to first appearances, there does not need to be anything practically indexical in such a computer's represen-

tation of time. In effect, a computer that worked in this way would have a system for generating stable files that refer to times, along with a labeling convention that enables the system to correctly store information about the present time. Call a computer that works in this way a *names-only* system.

A names-only system could be extended to have something like temporal indexicals. One way to do this would be to add a *now* buffer, much like the indexicals<sub>pr</sub>-basic system. Then information about the present time could be added simultaneously both to the *now* buffer and to the highest-valued stable file. This would create a hybrid system in which names and indexicals are equally basic. It is hard to see what advantage would be gained by this duplication of effort. But there is a more interesting, and perhaps more useful, possibility. For the sake of concreteness, imagine that the computer stores information about the weather. If a user enters **temperature(T)**, where **T** is the name of a time about which the computer has information stored – i.e., the “label” of a stable file – the computer will return the temperature information stored in that file. Now suppose the programmers want to add the possibility of reporting on the present weather, with a command like **temperature(now)**. It seems clear that this can be done without adding a *now* file. Instead, the computer should respond to **temperature(now)** with the temperature in the highest-valued file. On this way of developing the system, there would be no *now* file. Instead, **now** is acting as a pointer to whatever file happens to be the highest valued.

So far, I have restricted my attention to representations of times. But something similar seems possible for other sorts of representation. The crucial mechanism is the numerical labeling of files, which enables the system to select one file (the highest-valued) in a special way. In effect, the highest-valued file at a given time will have a special functional role. But one could imagine other means of giving files a special functional role. For example, a computer equipped with a GPS device could create stable files labelled with co-ordinates, and could store information about its present location in the file whose label matched the output of the GPS device. And of course there are means other than the file’s label; I see no reason why the subpersonal mech-

| Indexicals <sub>pr</sub> -basic System  | Names-only System   |
|---|---|
| <p>On each clock tick:</p> <ol style="list-style-type: none"> <li>1. Create a new stable, name-like file.</li> <li>2. Copy the contents of the NOW file into the new stable file.</li> <li>3. Erase the contents of the NOW file.</li> </ol> <p>To add information about the present time:</p> <ul style="list-style-type: none"> <li>• Update NOW file.</li> </ul> | <p>On each clock tick:</p> <ol style="list-style-type: none"> <li>1. Create a new stable, name-like file.</li> </ol> <p>To add information about the present time:</p> <ul style="list-style-type: none"> <li>• Update highest-valued stable file.</li> </ul> |

Figure 1: Comparison of the Indexicals<sub>pr</sub>-basic system and the Names-only system

anisms that link a being's perceptual input systems to files could not simply output to a stable file, rather than (as Recanati would have it) first routing them through a perceptual demonstrative file.

What is the upshot of this discussion for Recanati's views of mental files? First, I want to claim that:

**The Names Thesis** A names-only system is possible.

The Names Thesis does not claim that names-only systems involve no indexicality in Recanati's sense of that term. But they involve no practical indexicality – nothing that would ordinarily be thought of as context sensitivity. They are thus counterexamples to the Hierarchy Thesis. So the truth of the Names Thesis would undermine Recanati's picture of stable files, with their higher-order ER relations, as resting on a foundation of files based on first-order ER relations (many or most of which will be practically indexical).

Perhaps Recanati could give up this part of his picture. But the possibility of a names-only system raises a more troubling possibility for him: that *we* are names-only systems:

**The Names Question** Are we names-only systems?

A positive answer to the Names Question would undermine much of the interest of Recanati's view. If we are names-only systems, then we have only one type of file – the type associated with a higher-order ER relation – so Recanati's views about the significance of ER relations, indexicality, and so forth, would be more or less irrelevant to us. So what evidence do we have for the claim that we are not names-only systems?

Let me begin by setting one style of argument aside. Recanati frequently appeals to John Perry's view of mental files. Perry famously observed that identity claims involving both names and indexicals are very often – perhaps always – non-trivial and cognitively significant. For example:

- (1) That guy is Francois Recanati.
- (2) It is now ten o'clock.

(3) I am Derek Ball.

(In Recanati's terms, such claims typically do not involve co-reference *de jure*. See below for further discussion.) Does the non-trivial nature of such claims show that the file associated with "That guy" is of a different type than the file associated with "Francois Recanati", or that the file associated with "now" is of a different type than the file associated with "ten o'clock"? No. On Recanati's view, what is required for cognitive significance and non-triviality is (at least roughly) that the identity claim involve distinct *token* mental files. It does not require that these be tokens of distinct file types. So the Perry examples do not undermine the idea that all of our mental files are stable.

## 2 Conversion and Proper Function

I will return to the Names Question below; first, I want to examine a complication in Recanati's account. Consider again the names-only system that I described in the previous section. I claimed there that such files were from the beginning stable, name-like files. But this could be denied. In particular, on Recanati's view, files can *convert* from one type to another (2012, p. 81). In such conversion the "pile" of information in the file stays the same, but the ER relation associated with the file changes (2012, pp. 82-3). Recanati could claim that what I called the "names-only" system is really a system that (with each tick of the clock) generates a new (practically) indexical NOW file, and converts the previous NOW file into a stable file.

Recanati offers such an account in response to David Papineau's (2007) names-only description of a closely related case. When we see an object, Recanati claims, we are able to form a perceptual demonstrative file about that object. When we are no longer in a position to perceive the object, we should not retain a perceptual demonstrative file about it (since we will no longer stand in the relevant perceptual ER relation to the object). But, Recanati claims, we can convert this file into a memory demonstrative, which is based on an ER relation of memory. We thereby retain the information stored in the demonstrative file without violating the ER Norm.

| Names-Only Description   | Indexical <sub>pr</sub> -Conversion<br>Re-description  |
|--|--|
| <p data-bbox="331 468 576 495">On each clock tick:</p> <ol data-bbox="368 533 810 600" style="list-style-type: none"> <li data-bbox="368 533 810 600">1. Create a new stable, name-like file.</li> </ol> <p data-bbox="331 712 810 779">To add information about the present time:</p> <ul data-bbox="368 817 810 884" style="list-style-type: none"> <li data-bbox="368 817 810 884">• Update highest-valued name-like file.</li> </ul> | <p data-bbox="842 468 1086 495">On each clock tick:</p> <ol data-bbox="879 533 1321 667" style="list-style-type: none"> <li data-bbox="879 533 1321 600">1. Convert NOW file into a stable, name-like file.</li> <li data-bbox="879 638 1321 667">2. Create a new NOW file.</li> </ol> <p data-bbox="842 712 1321 779">To add information about the present time:</p> <ul data-bbox="879 817 1321 846" style="list-style-type: none"> <li data-bbox="879 817 1321 846">• Update NOW file.</li> </ul> |

Figure 2: Ways of describing the (alleged) names-only system

In short, there are two ways we can describe a case like my (alleged) names-only computer: in terms of stable files, or in terms of practically indexical files and conversion. What is at stake between the two descriptions? The first thing to note is that the descriptions do *not* differ about the sources from which the files can get information. On either story, the file can store information from any ER relation. The difference is in the *function* that the file is said to have. The indexical<sub>pr</sub> re-description claims that a particular ER relation is privileged: it is the relation that the file is based on, so that it is the function of the file to store information gained *by that relation*. The names-only description denies this; it claims that the function of the file is to store information gained by *any ER relation* that the file stands in to a certain object.

Now it may seem that this difference has consequences that will put us in a position to evaluate the two descriptions. Remember that violation of the ER Norm results in reference failure; or in other words, a file fails to refer if it fails to stand in the ER relation on which it is based to any object, or if it stands in this relation to more than one object. So it may seem that the two ways of

describing the scenario will have different consequences regarding when the file will fail to refer. According to the names-only description, the file will fail to refer if no object – or no *single* object – is the dominant source of the information in it. According to the indexical<sub>pr</sub>-conversion re-description, it fails to refer before the conversion if it fails to stand in the relevant indexical relation to any object, or if stands in that relation to more than one object. (What happens after the conversion will depend on what type of file it is converted to; I'll return to this issue shortly.)

Let's set up a case to try to drive the predictions of the two descriptions apart. (For the sake of simplicity, I'll use Papineau's perceptual demonstrative case as a model, though similar cases could be developed for reference to times.) I see a bird, and judge that it is grey. On the basis of my memory I come to believe that last Tuesday, that bird was eating fish. In fact, the bird I saw last Tuesday is not the bird I am now seeing. At the same time, a friend, whom I presuppose to be watching the same bird, tells me that that is a fulmar. In fact, she is watching a different bird. I combine the information obtained from three ER-relations into the file that I opened when I saw the bird.

Now it is not entirely clear what the referent of the file might be in this sort of case. But there is some intuitive pull towards the claim that the referent is the bird I see, and this is what the Indexical<sub>pr</sub>-Conversion Re-description would predict. So is the Names-Only description wrong? Only if the bird I see is not the dominant source of information in the file. As long as the bird I see is the dominant source of information in the file, then the Names-Only story will also predict that it is the referent. And in general, cases in which it is plausible that the file refers to the bird I see will be cases in which it is plausible that the bird I see is the dominant source of information in the file. So we need to develop a case in which the information that is dominant in the situation is not derived from the bird I see. Suppose I infer: "That is a fulmar. That ate fish. Therefore some fulmars eat fish." Now I am not relying on information gained perceptually from the bird I see; instead, I am relying on the information about other birds gained through memory and testimony.

What is the referent of the file associated with "that" in my inference?

Again, it is hard to determine; but – granting (as Recanati would insist) that the inference is valid – there is some attraction to the thought that the file associated with “that” fails to refer. So, victory for Names-Only? Not so fast. The Indexical<sub>pr</sub>-s-Conversion description can duplicate this result. The proponent of the Indexical<sub>pr</sub>-Conversion story should claim that our feeling that the file associated with “that” fails to refer shows that other ER-relations have become essentially involved in the file’s functioning; that is, just before the inference, the file has been converted from an indexical<sub>pr</sub> file to a stable file. In this case, the Indexical<sub>pr</sub>-Conversion story will have exactly the same consequences with respect to the reference of the file as the Names-Only story.

I want to draw two conclusions from this discussion:

1. In cases where the Indexical<sub>pr</sub>-Conversion story makes plausible predictions, the Names-Only story can duplicate them by appealing to the notion of a *dominant source of information*.
2. In cases where the Names-Only story makes plausible predictions, the Indexical<sub>pr</sub>-Conversion story can duplicate them by appealing to the notion of *conversion*.

These conclusions entail that it is very difficult to drive a wedge between the predictions of the two descriptions of the case. This difficulty generalises: it seems to me that almost any system could be described either in terms of stable files and dominance, or in terms of practically indexical files and conversion.<sup>1</sup> But this entails that it is very difficult to tell whether a given system involves indexical<sub>pr</sub> files or not. And in particular, it is hard to see what evidence there could be for Recanati’s view that we are indexical<sub>pr</sub>-basic systems rather than Names-only systems. So it is hard to see how Recanati can assume a negative answer to the Names Question.

---

<sup>1</sup>One problem is that Recanati tells us relatively little about what conversion involves. What does a mind have to do to convert one file type into another? How can one change the function of a file? Perhaps further development of the notion of conversion would put us in a better position to resolve this debate.

### 3 Co-reference and Transparency

The considerations of the previous two sections do not show beyond a doubt that there are no practically indexical files. But they do show that the phenomena can be described using Recanati's version of the file metaphor in a way that falsifies one of his main claims. In this section, I want to show that Recanati faces a similar problem with respect to another of his main claims: that sameness of mental files is *transparent*.

Recanati emphasises the importance of the phenomenon of *de jure* co-reference. The paradigm case of *de jure* co-reference, according to Recanati, is an anaphoric pronoun and its antecedent; for example, there is *de jure* co-reference in a sentence like, "Cicero<sub>i</sub> is bald and he<sub>i</sub> is fat", but (in the normal case) not in a sentence like, "Cicero is bald and Tully is fat". Moreover, Recanati claims, an analogous phenomenon can take place at the level of thought, and this can be explained in terms of mental files: two pieces of information are *de jure* co-referential if they "occur in the same file without the benefit of a prior judgement of identity or linking operation" (2012, p. 95).<sup>2</sup>

The distinction between *de jure* co-reference and non-*de jure* co-reference corresponds closely to a distinction between inferences that (in Campbell's phrase) "trade on" or presuppose identity, and those that assert identity. For example, it is natural to think that (*ceteris paribus* – Kripke puzzle cases and the like aside) one may validly infer from (4) and (5) to (6) without relying on further premises:

(4) Cicero is bald.

---

<sup>2</sup>Recanati also offers the following account of the phenomenon: "The basic criterion for *de jure* co-reference is, roughly, that anyone who wonders whether the two terms are co-referential (instead of taking for granted that they are) is someone who does not fully understand the utterance. In other words, *de jure* co-reference entails knowledge of co-reference on the part of the competent language users" (2012, p. 106, footnote omitted). In my view, Recanati's characterisation of the phenomena in terms of a connection between understanding and knowledge should be rejected for Williamsonian (2000, ch. 4, 2007, ch. 4) reasons. But it seems clear that there is a useful distinction between cases of co-reference due to anaphora on the one hand, and (say) co-reference between "Hesperus" and "Phosphorus" on the other, whatever the correct characterisation of that distinction ultimately may be.

- (5) Cicero is tall.  
 (6) Someone is tall and bald.

Indeed, the supposition that further premises asserting the identity of Cicero and Cicero are required seems to land one in an infinite regress. This contrasts with an inference from (4) and (7) to (6):

- (7) Tully is tall.

Such an inference seems (*ceteris paribus*) to require a further premise: that Cicero *is* Tully. Recanati cashes the distinction between the two inferences out in terms of *de jure* co-reference: in the first inference, the occurrences of “Cicero” in the two premises are *de jure* co-referential – the two occurrences of “Cicero” are associated with the same file, or the information that Cicero is bald and that Cicero is tall occur in the same file – while in the second inference, the occurrence of “Cicero” in the first premise is *not de jure* co-referential with the occurrence of “Tully” in the second premise – the two occurrences are associated with different files.

Recanati also endorses a closely related claim: that sameness of mental file is *transparent* in the following sense:

**The Transparency Thesis** “If it seems to you that two tokens [i.e., token mental representations, or token words as you are using them] ‘obviously and uncontrovertibly’ mean the same, then they do mean the same and co-refer (if they refer at all)” (Recanati (2012, pp. 117-8), citing Schroeter (2007)).

Recanati is not entirely explicit about how transparency is explained on his system, but I take it that the idea is something like this. Part of what it is for two pieces of information to be in the same mental file is that they be taken to concern the same object (2012, p. 101). So if you take two pieces of information to concern the same object, that will make it the case that the two pieces of information occur in the same mental file. So you cannot “mistake one conceptual content for another” (Schroeter (2007, p. 597),

quoted in Recanati (2012, p. 117)); if you take two pieces of information to be in the same file, they will be in the same file. Similarly, if you take it that two terms “obviously and uncontroversibly” mean the same and co-refer (as you are using them), then this will make it the case that you are associating them with the same file (and so they do co-refer, at least as you are using them).

Recanati mentions as potential problems for the transparency thesis what he calls “inverse Paderewski cases” (2012, p. 116): cases in which a thinker is acquainted with two bearers of the same name, and does not realise that they are distinct people. In such cases, Recanati claims, a thinker may have only a single mental file, in which she (mistakenly) stores information about both bearers of the name. This file, Recanati claims, fails to refer (at least in most circumstances –setting aside situations in which one or the other source may be dominant).

This description is plausible in simple cases. But the screw can be turned. One mark of rationality is the ability to recognise our mistakes. We may be accustomed to using a certain form of inference that we later come to regard as invalid. For example, I might be accustomed to inferring from premises of the form **All Fs are G** to conclusions of the form **There is at least one F**; but upon reflection, I may become convinced that “all” in English is not existentially committing, and so that these inferences are invalid (despite the fact that they once seemed obviously valid to me). Consider the following case (based on a true story):

**Recognised Equivocation** I judge:

- (8) Barry Smith is a philosopher of mind and language who works in London.

I then judge:

- (9) Barry Smith is a ontologist who works in Buffalo.

I infer from these premises (and no others) that someone is a philosopher of mind and language who works in Buffalo. It initially seems

obvious to me that the inference is valid. (I suppose that the philosopher in question must split his time between the two universities.) But upon further reflection, I decide that no single person could have distinguished himself in two fields in this way. I judge that my information must be about two distinct people who share the same name. I conclude that my initial inference was invalid. I judge:

(10) Barry Smith is not an ontologist

(intending my use of “Barry Smith” to be *de jure* co-referential with the first premise of my initial inference). I then infer (from (8) and (10)) that some philosopher of mind and language is not an ontologist. I also judge:

(11) Barry Smith does not work in London

(intending my use of “Barry Smith” to be *de jure* co-referential with the second premise of my initial inference). I infer (from (9) and (11)) that some ontologist does not work in London.

To sum up the situation:

1. At time  $t_1$ , I make an inference that trades on the identity of the referent of “Barry Smith” in (8) and the referent of “Barry Smith” in (9).
2. At time  $t_2$ , I make an inference that trades on the identity of the referent of “Barry Smith” in (8) and the referent of “Barry Smith” in (10), and another inference that trades on the identity of the referent of “Barry Smith” in (9) and the referent of “Barry Smith” in (11). I intend that the occurrences of “Barry Smith” in (10) and (11) do not co-refer, and I also deny that the occurrences in (8) and (9) co-refer.

The case produces two problems for Recanati’s view. The first is that it seems – at least by Recanati’s lights – that the occurrence of “Barry Smith” in (8) is *de jure* co-referential to the occurrence in (9), that the occurrence in (9) is *de jure* co-referential to the occurrence in (11), that the occurrence in

(10) is *de jure* co-referential to the occurrence in (8), but that the occurrence in (10) is *not de jure* co-referential to the occurrence in (11). That is, it seems to be a case in which the transitivity of *de jure* co-reference fails, and (as I will argue) fails in a way that is difficult for Recanati to accommodate. The second, and deeper, problem is that the case seems to be a counterexample to transparency. At the time I begin the inference, it seems obvious that the tokens of “Barry Smith” in (8) and (9) are co-referential. But I come to judge that I was mistaken, and that they in fact fail to co-refer. By my own lights, what once seemed to be an obviously valid inference turned out not to be. Let me consider each problem in turn.

Recanati gives an account of some related cases due to Angel Pinillos (2011). Pinillos claims (and Recanati agrees (2012, p. 110)) that in the sentence, “We were debating whether to investigate both Hesperus<sub>1</sub> and Phosphorus<sub>2</sub>; but when we got evidence of their true identity, we immediately sent probes there<sub>1,2</sub>”, both “Hesperus” and “Phosphorus” are *de jure* co-referential with “there”, but “Hesperus” is not *de jure* co-referential with “Phosphorus”. Recanati agrees that the example shows that *de jure* co-reference is not transitive, and that this entails that *de jure* co-reference cannot always be a matter of being associated with the same mental file. But Recanati claims that there is an alternative account available to the mental file theorist. In cases of what he calls “merging”, when one recognises that two distinct files refer to the same object, one will create a third file, copy all of the information from the initial files into the new file, and remove the initial files. Recanati describes Pinillos’s sentence as involving a case of “partial merging”: upon recognising the identity of Hesperus and Phosphorus, one creates a new file, copies the information from the two initial files into it, but – unlike in the case of full merging – one does not remove the initial files (2012, p. 111). This explains the case, since files that partially merge are *de jure* co-referential to the file that they merge into, while remaining not *de jure* co-referential to each other.

Recanati might attempt to describe the case of revealed equivocation in a similar way. In this case, the process would involve a sort of reverse merger: a branching. One would begin with a single file. Upon coming to believe that

the information in the file originated from more than one source, one would form two files, and copy information from the initial file into the two new files. This copying process might not be straightforward; one would have to disentangle which bits of information one had obtained from which source and put the information in the appropriate files accordingly (and perhaps there will be some pieces of information that one will not be able to place in either file, because one is not sure of their source.)

Now the analogue of Recanati's partial merger strategy would be to appeal to "partial branching". In this sort of case, one would form two new files, copying information as best one could, but would also retain the original file. Cases of this sort seem possible. But they would not help Recanati defuse the problem, since partial branching cannot result in co-reference in this sort of case. By Recanati's lights, the original file (as used in the inference described) fails to refer, since no single object is the dominant source of the information in it. But (barring any further problems) the two new files that result from the branching should refer. So the new files cannot co-refer with the original file.

Perhaps Recanati could add some operation on mental files to his arsenal in order to preserve the idea that there is *de jure* co-reference in all the right places in the case. But there remains the problem that the case poses for the transparency thesis. In short, the problem is this. At the time of my initial inference, I regarded the occurrences of "Barry Smith" in the two premises as obviously co-referential. But after further reflection and reasoning (which notably involves empirical information about what it is possible for one person to accomplish), I come to decide that I was wrong. According to my later judgement, what once seemed obvious turned out to be false; the two occurrences turned out not to be co-referential, and my inference turned out to be invalid. But since the transparency thesis has it that "If it seems to you that two tokens 'obviously and uncontroversibly' mean the same, then they do mean the same and co-refer (if they refer at all)" (2012, pp. 117-8), it is a consequence of the subject's own judgement about the case that transparency has failed in the first inference.

Recanati might try to argue that the subject's judgment is mistaken. The subject has correctly recognised that something has gone wrong with the inference, he might claim, but the diagnosis is wrong: in fact, the occurrences of "Barry Smith" are de jure co-referential, so that the inference is valid, but they both fail to refer.

The problem with this response is that it ignores the inferences that follow. After I judge that my initial inference was invalid, it seems obvious to me obvious that the occurrences of "Barry Smith" in (8) and (10), and in (9) and (11), are co-referential. Transparency entails that this seeming is accurate. But again, it is hard to see how this is compatible (given Recanati's commitments) with (i) the claim that the occurrences of "Barry Smith" in (8) and (9) fail to refer; and (ii) the extremely plausible claim that the occurrences of "Barry Smith" in (10) and (11) do not fail to refer. So it seems that Recanati could only maintain transparency between (8) and (9) in this way by letting transparency fail between (8) and (10) and between (9) and (11).

Alternatively, Recanati might try to argue that the case is not one in which the antecedent of the transparency principle is satisfied. After all, the antecedent stipulates that the two tokens must seem obviously *and incontrovertibly* to mean the same. But, Recanati might claim, in the present case, the two occurrences could not *incontrovertibly* seem to mean the same, since the appearance is quite quickly controverted. In effect, the suggestion would be that if two tokens seem incontrovertibly the same to a person, the person will not give up the idea that they are the same – at least not as a result of the sort of reasoning sketched in this case, and perhaps not under any circumstances.

At this point, it is an empirical question whether there are any cases in which two things seem "incontrovertibly" the same in the relevant sense. I am sceptical. The suggestion seems to underestimate our power of rational self-criticism. Quine, Putnam, Burge, Williamson, and others have made a powerful case that we can gain grounds for rational doubt of even those beliefs that seem obvious and trivial. The same is true of transitions in thought. Just as we can get evidence that our beliefs are false, we can get evidence

that our inferences are invalid; and to the extent that we are rational, we will respond to this evidence.<sup>3</sup>

#### 4 Conclusion

I have argued that Recanati's commitment to the indexicality of mental representation is not well-motivated, since his own theory allows for an apparently equally good description of the phenomena purely in terms of stable files. Moreover, I have shown that there are cases in which Recanati's account of co-reference *de jure* and transparency is difficult to maintain.

Regardless of the fate of his claims about indexicality and co-reference, Recanati has given us a number of subtle tools for describing mental representations. Perhaps not every distinction he introduces corresponds to a genuine difference. But there is no doubt that they should play a role in shaping future debate.

#### References

- Papineau, D. (2007). Phenomenal and perceptual concepts. In Alter, T. and Walter, S., editors, *Phenomenal Concepts and Phenomenal Knowledge: New Essays on Consciousness and Physicalism*, pages 111–144. Oxford University Press, New York.
- Pinillos, A. (2011). Coreference and meaning. *Philosophical Studies*, 154:301–324.
- Recanati, F. (2012). *Mental Files*. Oxford University Press, Oxford.
- Schroeter, L. (2007). The illusion of transparency. *Australasian Journal of Philosophy*, 85:597–618.

---

<sup>3</sup>Recanati writes, “a rational subject must be capable of reflecting critically upon his or her own thoughts; that sort of reflexive control over one’s thoughts is possible only if they are transparently accessible” (2012, p. 117). I accept the premise. But on my view, this premise supports exactly the opposite conclusion: if transparency is true, this limits our power to reflect critically on our thoughts.

Williamson, T. (2000). *Knowledge and its Limits*. Oxford University Press, Oxford.

Williamson, T. (2007). *The Philosophy of Philosophy*. Blackwell, Oxford.