



**HAL**  
open science

## Mental Files: an Introduction

Michael Murez, François Recanati

► **To cite this version:**

Michael Murez, François Recanati. Mental Files: an Introduction. *Review of Philosophy and Psychology*, 2016, 7 (2), pp.265 - 281. 10.1007/s13164-016-0314-3 . ijn\_01612895

**HAL Id: ijn\_01612895**

**[https://hal.science/ijn\\_01612895v1](https://hal.science/ijn_01612895v1)**

Submitted on 24 Aug 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# **Mental Files : an Introduction**

Michael Murez & François Recanati\*

In the first part of this editorial, we introduce to the topic of the special issue from the perspective of the three main disciplines that have appealed to the mental file construct: psychology, philosophy, and linguistics. In the second part we present the individual contributions to the issue, in connection with the various themes mentioned in the first part.

## **1. Mental files in psychology, philosophy, and linguistics**

### *1.1. Object files*

The notion of an *object file* is used in psychology to characterize visual representations of objects at an intermediate level between low level processing of sensory features and high-level placement of objects in conceptual categories or kinds. The object filing system supports the individuation and tracking of particulars, while allowing information about their features to be stored, updated, and retrieved. It has to decide when visually encountered elements should be counted as different stages of the same persisting object, and so should be assigned the same file, and when not.

A signature trait of the object filing system is that the features that matter to its operations are primarily spatiotemporal, rather than qualitative. Flombaum, Scholl, Santos (2009) dub the principle according to which, for object files, spatiotemporal factors take priority over featural similarity the ‘principle of spatiotemporal priority’. They argue that it constitutes a ‘fundamental principle of object persistence’ in human vision. A simple demonstration is the well-known ‘tunnel effect’ (Burke 1952). If the time it takes an object to disappear then reappear from behind an occluder is roughly the same as the time it would take an object to travel behind the occluder, and if the object’s way of moving is appropriate, viewers irresistibly experience a single persisting object, which is temporarily hidden behind the occluder. This is the case even when the reappearing object has very different surface qualities from the disappearing one – for example, a yellow ball can be perceived to ‘morph’ into a red cup. This experience as of one single persisting object is characteristically encapsulated, or resistant to influence from higher-level judgments of identity, which suggests that it is a perceptual effect.

---

\* Authors listed in alphabetical order.

One of the main sources of evidence for object-files are downstream attentional effects called ‘object specific preview benefits’ (OSPBs) (Kahneman, Treisman & Gibbs 1992, Noles, Scholl, & Mitroff 2005). In the object reviewing paradigm, subjects see a preview display in which visual information (a letter) is associated with ‘visual objects’ (square boxes). Subjects are faster at naming the letter when it reappears in the same box, after the letter has temporarily disappeared and the box has moved. Importantly, the design of the experiment ensures that the effect cannot be explained by the location in which the letter has reappeared. Intuitively, the priming-like effect ‘travels’ with the moving box (the visual object). This suggests that *objects per se* serve as units for the allocation of visual attention. OSPBs occur when (time-slices of) objects are assigned the same object file, ‘correspondence’ between them having been computed on the basis of predominantly spatiotemporal factors, enabling faster access and retrieval of the features already stored in the file.

The principle of spatiotemporal priority operates in a wide variety of cases. Especially striking in this respect are effects observed not in experiments on adult vision, but in the initially independent research tradition which studies infants ‘object concept’ (Spelke 1990, Baillargeon 1995, Carey 2009). As was discovered, “[t]here are intriguing parallels between research on the infant's object concept and research in the nature of object-based visuospatial attention in adults” (Scholl and Leslie 1998: 66). Infants innately expect objects to obey certain spatiotemporal principles, which together define the notion of a *Spelke object* (roughly, objects which move as bounded, coherent wholes) as opposed (e.g.) to ‘stuffs’ (like water or sand) which do not possess clear boundaries or shift shapes as they move (cf. Rips and Hespos 2015). Carey and Xu (2001) emphasize that infants strongly privilege spatiotemporal information over property and kind information in individuating and tracking objects, as measured using habituation/dishabituation paradigms (infants look longer when they are surprised by something unexpected). In one experiment, ten month olds see two stimuli appear successively from behind each side of an occluder, which differ in kind or properties (e.g., a red and a blue ball). Adults would draw on their conceptual resources (objects don’t suddenly change properties and kinds) to judge that there are two different objects, which become successively visible. But ten month olds expect one object behind the occluder (though it can be shown independently that they are aware of the change in properties). Thus, infants initially prioritize spatiotemporal information in computing object persistence, much as adult vision does in, e.g. the tunnel effect. Carey and Xu (2001: 210)

conclude that “young infants' object representations are the same natural kind as the object files of mid-level vision” (see also Scholl 2001, Carey 2009 ).

In addition to the convergence between these two trends of research (on the infant's object concept and on adult object-based visual attention), there is another convergence that is worth noticing. Pylyshyn has emphasized the brute causal nature of the determinants of reference for object files (Pylyshyn 2000, 2001, 2007, Fodor & Pylyshyn 2014). His FINSTs are indexing devices which trigger the opening of an object file about the indexed object. FINSTs are responsive to features in the visual landscape: by exhibiting the relevant features the appearance of an object in the visual environment may ‘grab’ an index in such a way that an object file is opened for that object. But *the features responsible for index grabbing are not themselves represented in the content of the file which is thereby opened*. The mechanism that fixes reference is entirely ‘preconceptual’, Pylyshyn says:

With the typical sort of conceptual representation, there is no way to pick out an individual in the world other than by finding the tokens in a scene that fall under a particular concept, or satisfy a particular description, or that have the properties encoded in the representation. What I will try to show is that this cannot be what goes on in general; it can't be the case that the visual system can only pick out things in the scene by finding instances that satisfy its conceptual representation. There are phenomena that suggest that the visual system must be able to pick out individuals in a more direct manner, *without using encoded properties or categories*. If this claim is correct then the visual system needs a mechanism for selecting and keeping track of individual visual objects that is more like a demonstrative reference (the sort of reference we make in language when we use demonstrative terms like *this* or *that*) than a description. And that, I suggest, is why we must have something like a visual indexing mechanism which preconceptually picks out a small number of individuals, keeps track of them, and provides a means by which the cognitive system can further examine them in order to encode their properties, to move focal attention to them or to carry out a motor command in relation to them. (Pylyshyn 2001 : 130 ; emphasis ours)

As Pylyshyn notes, this proposal is very similar to an idea which is ubiquitous in the philosophical literature on singular thought, namely, the need for *non-descriptive grounding* for thoughts about particulars. Mental files, as characterized in the philosophical literature, are

devices of direct reference whose deployment makes it possible to entertain *singular thoughts*, i.e. thoughts that are about particular objects *rather than* about whatever possesses certain features or satisfies such and such description. To this issue we now turn.

### *1.2. Mental files as singular modes of presentation*

One leitmotif in the philosophy of language and mind of the past fifty years has been its *anti-descriptivism*. Some objects are represented descriptively, via their qualitative features which are *themselves* represented; but not all objects can be represented in this indirect manner (via the representation of their properties). According to a well-known argument known as the buck-passing argument,

There must be some representations whose referential properties are not parasitic on those of others, else language as a whole is cut loose from the world. Description theories pass the referential buck, but the buck must stop somewhere. It stops with theories (...) that explain reference in terms of direct relations to reality (Devitt 2014 : 477).

The paradigm of direct reference to particulars is the *demonstrative* paradigm. In a perceptual encounter with an object, the subject can think about the object demonstratively ('that thing'), whether or not he or she is able to correctly classify it as belonging to a given kind. Demonstratives (whether in language or thought) are such that their reference is determined 'relationally', not 'satisfactionally' (Bach 1987). Recanati (1993: 171-72) uses Treisman's object files ('it's a bird, it's a plane, it's Superman') to illustrate the functioning of demonstrative representations of particulars in thought. Such representations he construes, following Perry (1980), as mental files whose reference does not depend on the properties or category information to be found in the file but on tracking relations to the object which trigger the opening of the file. Pylyshyn similarly stresses the analogy between his FINSTs and demonstrative reference as characterized in the philosophical literature :

Although it is not often recognized we do, under certain conditions, represent some things without representing them in terms of concepts. We can refer to some things, as I will say, preconceptually. For example, in the presence of a visual stimulus, we can

think thoughts such as 'that is red' where the term 'that' refers to something we have picked out in our field of view without reference to what category it falls under or what properties it may have. A term such as this or that is called a 'demonstrative'. Philosophers like Perry (1979) have argued that demonstratives are ineliminable in language and thought. The reasons for the ineliminability of demonstratives in language and thought also apply to visual representations. Not only can we represent visual scenes in which parts are not classified according to some category, but there are good reasons why at least some things must be referenced in this preconceptual way. If we could only refer to things in terms of their category membership, our concepts would always be related only to other concepts (the concepts for categories) and would never be grounded in experience. Sooner or later the regress of specifying concepts in terms of other concepts has to bottom out. Traditionally, the 'bottoming out' was assumed to occur at sensory properties, but this 'sense data' view of concepts has never been able to account for the grounding of anything more than simple sensory concepts and has been largely abandoned. The present proposal is that the grounding begins at the point where something is picked out directly by a mechanism that works like a demonstrative. (Pylyshyn 2001 : 129)

This is the buck-passing argument again. Pylyshyn formulates it in terms of the conceptual/preconceptual distinction, but that distinction suffers from an ambiguity which is worth dispelling at this point.

FINSTs and the object files they initiate provide grounding that is preconceptual in the sense that it is a 'brute causal mechanism' occurring in early vision (Pylyshyn 2001 : 147). But demonstratives as they occur *in thought* are bound to be conceptual, in the minimal sense that they are thought constituents. Pylyshyn insists that even they are 'preconceptual', because the reference-fixing mechanism is independent of any description used by the subject to categorize the object. This, we believe, confuses *descriptive* and *conceptual*. The point is not merely terminological. What matters is that *we need two distinctions*, not merely one. The first distinction we need is that between the reference mechanisms at work in vision and 'input systems' more generally (Fodor 1983), and the reference mechanisms occurring in thought. Admittedly, there is a strong analogy between them, an analogy which Pylyshyn and Recanati both emphasize. The analogy suggests that object files are the *precursors* of the mental files which philosophers talk about in connection with singular thought. Similar though they are to object files, however, the mental files posited by philosophers are bound to

be conceptual *because* they are thought constituents. Qua thought constituents, they participate in cognitive practices regulated by rationality principles, and they involve 'representing-as'.

In the philosophical literature, it is emphasized that *not all concepts are descriptive*. Demonstrative concepts such as 'that thing' are nondescriptive (i.e. their reference is fixed relationally rather than satisfactionally). The descriptive/nondescriptive distinction applies *within* the conceptual realm. In contrast, Pylyshyn tends to equate the two distinctions (conceptual/nonconceptual, and descriptive/nondescriptive), to the point of denying that demonstrative representations, with their nondescriptive mode of reference-fixing, are instances of 'representing as'. If something is represented *conceptually*, Pylyshyn says, it is *represented as* something, but FINSTs and object files represent in a more direct manner. They are preconceptual in the sense of *not* involving representing-as:

The indexing system latches on to certain kinds of spatiotemporal objects because it is 'wired' to do so, or because it is in the nature of its functional architecture to do so, not because those entities satisfy a certain cognitive predicate — i.e. not because they fall under a certain concept. This sort of causal connection between a perceptual system and an object in a scene is quite different from a representational or intentional or conceptual connection. For one thing *there can be no question of the object being misrepresented since it is not represented as something*. (Pylyshyn 2001 : 147 ; emphasis ours)

It is true that Pylyshyn's visual indexes, and the object files they initiate, are nonconceptual because of the early level of processing they belong to. But demonstrative representations *in thought*, which Pylyshyn describes as *also* based on a direct ('preconceptual') connection to the world, undoubtedly represent their reference 'as something'. Misrepresentation is *always* possible in thought, even if the object is represented demonstratively. Thus in a well-known example due to David Kaplan (1989), the subject looks at himself in the mirror, without recognizing himself, and thinks, about the person seen : 'His pants are on fire'. The subject does not act in the way he would if he thought 'My pants are on fire'. The connection to action (which Pylyshyn 2000 emphasizes, following Perry), and the fact that the subject assents to one thought ('his pants are on fire') but not to the other ('my pants are on fire'), suggest that they are distinct thoughts, with distinct cognitive contents, despite the fact that the first person and the demonstrative refer to the same

individual. What distinguishes the two thoughts is the *mode of presentation* of the individual they are both about.

The involvement of modes of presentation corresponds to the idea of representing-as and is the hallmark of conceptual thought, according to Pylyshyn : modes of presentation provide what he calls ‘a representational or intentional or conceptual connection’ to the thing thought about. Now, in contrast to the psychologist’s object files, the mental files posited by philosophers are essentially meant to play the mode of presentation role (Recanati 2012, 2016). This role is characterized through *rationality constraints* such as the following :

### **Frege’s Constraint**

If  $m$  is a mode of presentation under which a minimally rational person  $x$  believes a thing  $y$  to be  $F$ , then it is not the case that  $x$  also believes  $y$  not to be  $F$  under  $m$ . In other words, if  $x$  believes  $y$  to be  $F$  and also believes  $y$  not to be  $F$ , then there are distinct modes of presentation  $m$  and  $m'$  such that  $x$  believes  $y$  to be  $F$  under  $m$  and disbelieves  $y$  to be  $F$  under  $m'$ . (Schiffer 1978 : 180)

### **Trading-on-Identity Constraint**

From ‘ $a$  is  $F$ ’ and ‘ $b$  is  $G$ ’, one can rationally infer that some  $x$  is  $F \& G$  only if ‘ $a$ ’ and ‘ $b$ ’ are associated with the same mode of presentation (not merely the same reference). If ‘ $a$ ’ and ‘ $b$ ’ are coreferential without being associated with the same mode of presentation, an additional premiss ‘ $a = b$ ’ is required to support the inference to ‘ $\exists x Fx \& Gx$ ’.  
(Campbell 1987 : 275-78)

Let us assume that mental files are, indeed, what plays the mode of presentation role (the ‘realizers’ of the role). Frege thought of modes of presentation as descriptive/satisfactional, and there are theorists who construe mental files themselves in this manner. Thus Grice, an early mental file theorist, characterized ‘dossiers’ as collections of descriptions presupposed to corefer (Grice 1969). But the anti-descriptivist movement in the seventies established that the descriptions a subject can provide are not what fixes the reference of her singular terms. Causal relations crucially come into play. In the case of indexical thoughts involving demonstrative files or first person files, as in Kaplan’s example, what fixes the reference are token-reflexive relations to the file. Demonstrative and first person files are nondescriptive because their reference is fixed via acquaintance relations



rather than satisfactorily, yet they are bona fide concepts. The subject who thinks ‘My pants are on fire’ represents himself (the possessor of the pants) in the first person way (via the self file). This is *representing as*. And Kaplan’s example is an example of *misrepresentation*. The subject who thinks ‘His pants are on fire’ misrepresents himself as someone else (through a demonstrative file). Indexical concepts are, indeed, concepts, despite the fact that their reference depends on nonconceptual, contextual relations.

The same thing holds for the singular concepts (mental files) associated with names such as *Hesperus* or *Phosphorus*. These files too are based on an array of acquaintance relations, even if the notion of ‘acquaintance’ has to be extended to take account of testimony as a source of first order information. And they are bona fide concepts, supporting representing-as: Venus can be thought of alternatively as ‘Hesperus’ (the Evening Star) or as ‘Phosphorus’ (the Morning Star), just as Paderewski can be thought of alternatively as Paderewski the musician or Paderewski the politician, and Cicero as ‘Cicero’ or as ‘Marcus Tullius’ (not to mention Orcutt, Lingens, and other legendary figures).

### *1.3. Mental files and ‘discourse referents’*

In philosophy of language and linguistics, the mental file story initiated a new, dynamic perspective on linguistic meaning. Strawson, one of the early proponents of mental files, emphasized the role of *file management* in linguistic communication. Understanding an utterance consists in incorporating or integrating the information it carries into the subject’s worldview or mental encyclopedia, consisting of a network of mental files (Strawson 1971, 1974). To do so the subject must either select the relevant file into which the information provided by the utterance must be fed, or, if there is no such file, to open a new one. Which operation is appropriate upon receipt of linguistic information is indicated by linguistic devices whose function is to guide understanding (Chafe 1976, 1994) :

- The role of *definite markers* is to indicate that the listener is expected to activate a *preexisting* file ; in contrast, indefinites instruct the listener to open a *new* file (Hawkins 1978, Heim 1983).

- *Information structure* (so-called ‘packaging’) determines which of the preexisting files activated by a given utterance is primarily concerned by the information it conveys and serves

as ‘topic’ for the utterance (Reinhart 1981, Erteschik-shir 1997, Lambrecht 1996, Vallduvi 1992).

- The *Givenness Hierarchy* maps the type of referring expression used to the cognitive status of the file it concerns. Unstressed pronouns are associated with information to be fed into an *already activated* file, while e.g. definite descriptions require (re)activation of a file that preexists in long term memory but is not already active in short term memory (Prince 1981, Ariel 1988, Gundel, Hedberg & Zacharski 1993).

Another kind of file-managing instruction is provided by *identity statements*. Identity statements such as ‘Hesperus is Phosphorus’ were invoked by Frege to establish the need for a level of ‘mode of presentation’ (or ‘sense’) in addition to reference. ‘Hesperus is Phosphorus’ is informative, Frege says, in contrast to ‘Hesperus is Hesperus’ or ‘Hesperus is itself’. It is informative because the modes of presentation associated with the names ‘Hesperus’ and ‘Phosphorus’ are distinct (so it is not known a priori that they corefer). In the mental file framework that means that distinct files are associated with the names ‘Hesperus’ and ‘Phosphorus’. According to Strawson, however, the information conveyed by ‘Hesperus is Phosphorus’ is not first order information to be fed into one or both of these files, but higher order information pertaining to file management : the listener is instructed to *merge* the Hesperus file and the Phosphorus file, instead of retaining two separate files (Lockwood 1971, Strawson 1974).

Let us turn to the centerpiece of dynamic semantics: the notion of a discourse referent (Karttunen 1976, Kamp 1981). A discourse referent is like a referential ‘peg’ with associated attributive features that are updated as discourse proceeds. That is very much like a mental file. Mental files consist of three components : the file itself as a mental particular, the (reference-fixing) relations it is based on, and the informational content of the file. Discourse referents also are particulars endowed with numerical identity (there can be two distinct discourse referents with the same associated information), and, at any given point in the discourse, they carry informational content (viz. the set of associated conditions). What about the third component, the reference-fixing relations ? Here we find the major difference between mental-file-like objects in linguistics and their counterparts in philosophy.

Mental files are ‘anchored’, at least internally. They are about the object (if any) they are suitably related to. But in linguistics there are two kinds of discourse referents : those that,

like mental files, are internally anchored, i.e. which are supposed to refer in the actual world through relations to objects in the environment, and those that are *not* anchored and are *not* supposed to refer in the actual world (Kamp 1990, 2015). Anchored entity representations, as Kamp calls them, are mental files in the standard sense, but unanchored entity representations are very different from mental files as we have characterized them so far. They lack the third component – the reference-fixing relations. They are like the *descriptive* files discussed in Goodman’s paper in this issue.

Dynamic semantics initially started by exploiting the definiteness criterion. If indefinites open mental files and definites activate preexisting mental files, then we have to say that in ‘I dreamt that there was *a new President* and that *he* called me on the phone’ a file for the new President is opened by the indefinite in the sentential complement of the verb in the first clause ; that file is associated with the definite ‘he’ in the second clause, in such a way that the predicate  $\lambda x. x \text{ called me on the phone}$  is fed into the file. Such files are temporary descriptive files which, in the course of linguistic processing, are used to build up a model of the situation talked about (Karttunen 1976). Of course, the new President *only* exists in the dream. The speaker/thinker is not committed to its reality. In a case like this, the (unanchored) discourse referent for the new President is not present in the domain of discourse representation structures at the top level — the level which corresponds to the speaker’s beliefs about reality. The discourse referent for the new President is only present at a subordinate level, in the discourse representation structure representing the content of the dream. Discourse referents endowed with that ancillary function (representing the content of suppositions or of mental or linguistic acts or similar things) do not refer and are not supposed to refer. They merely simulate reference.<sup>1</sup>

With such discourse referents, we have gone a long way from object files ! One of the central issues that arise in the mental file framework is that of the relations between unanchored files, especially the temporary/subordinate variety, and their standard referential

---

<sup>1</sup> Simulation is what can be observed in the closely related phenomenon of *arbitrary reference*. If I say ‘let  $n$  be a number between 0 and 9’ and start ascribing certain properties to  $n$ , I am not referring to an ‘arbitrary number’ distinct from all particular numbers (Fine 1983), nor am I referring to one particular number while being unaware of its identity (Breckenridge and Magidor 2012); rather, I *pretend* to refer to a particular number (in the shallow sense of ‘pretense’ glossed in Crimmins 1998). Recanati (2012, 2013) conjectures that all descriptive files simulate reference and are therefore parasitic on standard referential files.

cousins. An attractive view is the following (Recanati 2013). Just as object files may be construed as the precursors of standard referential mental files, the latter may be construed of the precursors of the nonreferential type exploited in discourse and linguistically driven thought. If the nonreferential use of mental files illustrated by ordinary discourse and thought rests on a form of simulation, then the primary use is bound to be the referential use, which the nonreferential use simulates.

## 2. Preview

This issue is divided in five sections. The first section is devoted to the crucial contrast between descriptive and non-descriptive reference fixing. The second section deals with issues of identity and confusion. In the third section, we have gathered criticisms of the mental file framework as standardly understood. The fourth and fifth sections are devoted to the role of files in language processing (section 4) and in the development of mindreading and related abilities (section 5). Each section contains two or three papers.

### 2.1 *Descriptive and non-descriptive reference fixing*

In ‘Object files, properties and perceptual content’, Santiago Echeverri argues that even in the case of *object files*, where reference-fixing is done preconceptually, what fixes reference is not a brute causal relation (as Pylyshyn holds) but the intentional content of the underlying perception. Without basic perceptual operations like grouping and figure/ground segregation, no visual index could be assigned and no object file opened, Echeverri points out ; but these perceptual operations determines *correctness conditions* and the representations they output can *misrepresent*. As he puts it, ‘the existence of these correctness conditions indicates that perceptual individuation is a primitive form of representing-as’. It follows that we should make room for *preconceptual intentional content* and allow that it (rather than brute causal factors) fixes reference and thereby makes it possible for conceptual intentional content to be singular and concern a particular object.

In ‘Mental Graphs’, Jim Pryor offers a graph-theoretic representation of the mental encyclopedia, with vertices (or nodes) representing objects and edges representing relations. This elaboration of the mental file framework has several advantages over standard versions. One advantage is the easy representation of *relational* information, which need not be encoded within the files but is now represented through the edges which link the files. (This

generalizes to all relations the manner in which identity is represented in the mental file framework — through linking.) Second and most important in the context of the section, the framework allows *distinct but indiscernible files*, i.e. files of the same type having the same referent and *containing the same information* (yet contributing differentially to content). In the neo-Fregean framework that corresponds to the idea that senses can differ *solo numero*. This is an extreme form of singularism, which normally goes together with what Pryor calls a ‘symbolist’ interpretation, i.e. the idea that *the numerical identity of the mental vehicles* can affect content. Pryor, however, rejects the symbolist interpretation, which invokes particular mental symbols (the files, or words in a language of thought). Graph-theoretic structure is something more abstract, he says. The difference between the two approaches is that one approach has to make sense of permutation possibilities which the other one can safely ignore :

If Alice really does have two numerically distinct but qualitatively indiscernible presentations of Bob, and these presentations are reasonably thought of as two *particular senses*, then mightn’t it make sense to talk about *two* (or more) situations in which these particular senses are permuted ? I don’t think that *ought* to make sense. There does not have to be anything in the mental reality we are modeling — at the level of abstraction we’re modeling it, which may be shared between Alice and other persons who realize the same structure in concretely different ways — that captures such differences. At this level of abstraction, the permutations aren’t meaningful. Properly understood, that’s the result a graph-theoretic framework for talking about Alice’s mind should deliver. (...) If we talk about *particular senses*, though, I worry that these permutational differences might be genuine possibilities.

Aidan Gray’s paper is concerned with the fundamental issue, whether or not we should make room for distinct but indiscernible files. In contrast to Pryor, he argues that we shouldn’t. He focuses on the files associated with proper names. Could a subject have two numerically distinct files associated with the same name type (e.g. ‘Paderewski’), referring to the same individual *and containing the same information* ? Gray says no. He defends the view he dubs ‘Minimal Descriptivism’ (the title of his paper) : if a competent speaker finds ‘a = b’ informative (where ‘a’ and ‘b’ are two token names), then *she must associate some information with ‘a’ which she does not associate with ‘b’*. Minimal Descriptivism, Gray argues, is compatible with the rejection of satisfactionalism. The post-Kripkean orthodoxy

rejects both satisfactionalism and Minimal Descriptivism, but this is a mistake due to an improper appreciation of the type of information processing name-based files involve.

Gray argues that the acquaintance relations appealed to by the file theorist must be conducive to *knowledge*. In this he follows Williamson :

A causal relation to an object (property, relation,...) is a channel for reference to it only if it is a channel for the acquisition of knowledge about the object (property, relation). (Williamson 2007, ch. 8, cited in Gray's paper)

But to be conducive to knowledge the relation to the object established through the name has to be such that the speaker *knows what she says* when she uses the name. That knowledge requirement constrains the files associated with the name : unless they accord with Minimal Descriptivism, the speaker fails to meet the knowledge requirement and lacks competence with the name. That is how Gray would presumably respond to Pryor's examples of distinct but indiscernible files : if they are associated with names, the files are defective — they cannot support competent use of the name.<sup>2</sup>

## 2.2 Identity and confusion

The paper by Perner and Leahy might have appeared in this section, for it bears on dual naming tasks and the representation of guises. (It might have appeared also in the previous section, since it bears on the role of sortal information in reference fixing.) Since we preferred to include it in section 5, on the development of mindreading and related abilities, the current section only contains two papers: one by Filipe Drapeau-Contim's, 'Mental Files and Non-Transitive De Jure Coreference', and one by Sam Wilkinson, 'A Mental File Approach to Delusional Misidentification'.

Coreference *de jure* — the topic of Drapeau-Contim's paper — is the relation that holds between two representations when they 'represent as the same' (Fine 2007), i.e. when referential divergence between them is ruled out on an a priori basis. In the mental file framework it is often assumed that coreference *de jure* is a matter of identity between the associated files, but certain examples due to Pinillos (2011) suggest that coreference *de jure* is

---

<sup>2</sup> As Gray acknowledges, this argument for Minimal Descriptivism does not generalize. It works only for names and does not apply for instance to purely perceptual examples of distinct but indiscernible files.

not transitive. These examples invoke fusion and fission of files and also vicarious files ‘indexed’ to other agents (a topic addressed in Perner’s and Leahy’s paper). Fusion of files corresponds to certain cases of confusion, but also to the use of ‘slash-terms’ such as *Hesperus/Phosphorus*. Drapeau Contim reanalyses the examples and argues that paradoxical consequences follow from the assumption that the slash-term ‘Hesperus/Phosphorus’ *de jure* corefers with each of the two basic terms ‘Hesperus’ and ‘Phosphorus’. Giving up that assumption enables Drapeau-Contim to preserve the thesis that coreference *de jure* is transitive, appearances notwithstanding.

In ‘A Mental File Approach to Delusional Misidentification’, Sam Wilkinson argues that delusions of identity such as the Capgras or the Fregoli syndroms result from a mismanagement of files. A Capgras patient is one who generates too many files and fails to ‘represent as the same’. A Fregoli patient suffers from the opposite deficit : they over-‘represent as the same’. These syndromes exploit the separation of the file from the predicative information in the file, only the file playing an individuative role. In the singularist framework in general, and the mental file framework in particular, individuation and description *can* come apart; in the Capgras and Fregoli delusions they *do* come apart. As Wilkinson writes,

The notion that someone could judge that someone is not S in spite of looking just like what the subject took S to look like (Capgras) or that someone could judge that someone is S in spite of looking nothing like them (Fregoli) is much less outlandish to the singularist than to the descriptivist.

### 2.3 Critics of the mental-file framework

Mental files involve three components, we said : the file itself (its numerical identity), the content of the file, and the acquaintance relations it is based on. The critics of the mental file framework gathered in this section each target one particular component and argue that it should be dissociated from the rest, thereby altering the framework in a significant manner.

It is the ‘informational content’ component which, in ‘Mental Files : What For ?’, Alfonso Losada proposes to dismiss. He contrasts the mental file view with the ‘austere view of the mind’, according to which the mind hosts attitudes towards articulated representations, where the articulated representations are composed of concepts. Some concepts are singular,

and the thoughts they are part of are singular thoughts. Losada's preferred version of the austere view is singularist, like the mental file view, but Losada objects to burdening the singular concepts with the task of 'storing' information pertaining to the referent of the concept. This additional task is unnecessary, he argues. If we opt for the austere picture rather than the mental file picture, the alleged informational content of the file is distributed over all the subject's attitudes involving the singular concept.

Losada raises the issue of *relational* information : how could the information that John loves Mary be included in the file for John given that that piece of information is (also) about Mary? Does that mean that the file for John, which hosts that information, would somehow contain the file for Mary? Of course, the file for Mary will also contain information involving John. Does it not follow that in such a situation a file will contain itself ? (The same sort of worry is expressed in Millikan 2012; see also Woodfield 1991 : 549.) Losada mentions Recanati's solution in terms of pointers, but it is worth stressing that Pryor graph-theoretic representation provides an elaboration of the mental file framework which is immune to Losada's criticism. In Pryor's framework, the information in a 'file' is indeed represented by the place of the corresponding vertex in the graph-theoretic network ; it is distributed over the subject's entire belief system.

The two other papers in the section target the acquaintance relation component of mental files. Carsten Hansen and Georges Rey focus on Recanati's book *Mental Files* and criticize the 'actualist' assumption it inherits from the discussion of singular thought since Evans. What fixes reference in Recanati's theory are what he calls 'epistemically rewarding (ER) relations' – relations of acquaintance on which the files are based. But ER relations are actual world relations : relations the file (qua mental particular) bears to the referent of the file in the environment. That entails that a file can only refer, and satisfactorily contribute to content, if the referent exists in the actual world. Hansen and Rey object that very often the referent targeted by a singular thought does not actually exist or, if it exists, cannot stand in causal-informational relations to the file. Rather than try to accommodate the counterexamples on a case by case basis, Hansen and Rey suggest that the mental file theorist would do well to get rid of the third component altogether (the acquaintance relations, with their actualist consequences), and to construe singular thoughts as a specific kind of 'internally focused computational state'. Kamp's distinction between internal and external anchors for discourse referents or entity representations is a first step in that direction (see



Maier's paper in section 4). Internal anchors correspond to Recanati's idea of normative presuppositions attaching to the files, presuppositions which may or may not be satisfied. On Recanati's view files are *supposed* to refer to objects in the environment through ER relations, but there are many cases in which they don't actually succeed in doing so, or only pretend to do so (simulation cases). This introduces a hierarchy, with some objects of thought being evolutionarily more basic than others : the basic cases are the true relational cases, on Recanati's account. Hansen and Rey object that Spelke-objects, which are elevated to the status of paradigm in the standard mental file story, are *not* representative of the whole class of objects of thought and should not be so elevated. They draw attention to 'the pervasiveness of the problem of non-Spelke objects even in elementary perception'. The things we think about are 'things' for us since we think about them, but that shouldn't commit *the theorist* to the actual existence of these things, even in a restricted range of cases serving as paradigm. 'It could turn out to be an important psychosemantic fact that humans *think* singular thoughts without there actually existing *any* individual objects that satisfy them'.

Rachel Goodman's paper, 'Against the Mental Files Conception of Singular Thought', also targets the third component (the acquaintance relations). She argues for dissociating the mental file construct and the idea of singular thought. If singular thought is a mode of reference fixing, namely *relational* reference fixing (reference fixing through acquaintance), then it must be conceded that some files are *not* singular because their reference is fixed satisfactorily. Not only is the reference of descriptive files fixed in this way ; the reference-fixing description plays various additional roles pertaining to file management — it *governs* the file.

In a case with these features, the marks of file-hood are in place but the reference determination for the file, the information marshaling strategy of the file, the possibilities for error and misinformation in the file, and the identity conditions of the file over time are determined *satisfactorily* rather than *non-satisfactorily*. Given our understanding of descriptive thought as satisfactory, and singular thought as non-satisfactory, we have reason to say that the content of thoughts employing such a file is descriptive. This is perfectly consistent with the file *being* a file.

Again, the mental file theorist could reply that descriptive files are evolutionarily less basic than regular relational files, and presuppose them. But their existence can hardly be doubted,

and any theory has to account for them. (Earlier in this introduction we pointed out that unanchored discourse referents in linguistics are functionally similar to descriptive files.)

#### *2.4 Mental files in language processing*

In ‘Mental Files and the Lexicon’, Luca Gasparri deals with the meaning of words. The body of encyclopedic information associated with a lexical entry is best understood as a mental file. Singular terms are associated with individual files, but common nouns, verbs etc. are associated with files for kinds, eventualities, properties, locations, etc. These files contain pieces of information which (on the standard contextualist picture) are recruited to build up the ad hoc concept which, on a particular use, the word contributes to the context-sensitive content of the utterance (Sperber and Wilson 1998, Recanati 2004). But the word is also associated with ‘proprietary information’ about the word itself and its phonological, syntactic and semantic properties. Gasparri conjectures that words are represented through *lexical files* in the mind of language users. Lexical files are mental representations of words, which are used both to detect the word when it is instantiated in discourse and to come up with a proper interpretation of its meaning in context. Crucially, a lexical file contains a pointer to the associated encyclopedic file. ‘The interpretation of lexical items is normally the result of context-sensitive processes whereby underdetermined semantic schemata are relativised to conceptual structures that exceed the boundaries of linguistic knowledge’. The conceptual structures are provided by the encyclopedic file associated with the word. So, to reformulate Gasparri’s proposal, the meaning of a word consists of three components : a recognitional component whereby the lexical files makes lexical access possible, syntactic and semantic schemata determining how the word composes with other words in the sentence, and a meaning-determining content. ‘The meaning determining content associated with words, qua psychological types, consists primarily of a bundle of instruction-like pointers that supervise the recruitment of the cognitive resources (analogical reasoning, world knowledge) needed to flesh out the semantic value of word occurrences in contexts’.

The other paper in the section is an overview, by Emar Maier, of the representation of singular attitudes in current versions of Discourse Representation Theory (Maier 2015, Kamp 2015). Maier stresses the analogies between the mental file framework and the way singular reference is handled in recent versions of DRT (an analogy already emphasized in Pryor’s paper). He also reminds us that Kamp always had a cognitive interpretation of the framework

in mind. The natural conclusion to draw is that DRT might be an appropriate formal framework for mental file theory.

Kamp and Maier view the attitudes (beliefs, desires, imaginations, etc.) as consisting of two components : a label (BEL, DES, IMG...) corresponding to the 'mode' of the attitude and a DRS corresponding to its content. A DRS, or discourse representation structure, itself has two components : a set of discourse referents (what the attitude is about) and a set of descriptive conditions on the discourse referents. The DRS that is, or represents, the content of an attitude may involve 'free' discourse referents, which need to be contextually bound. Such a DRS only has content relative to a context that provides values for the free discourse referents. The free discourse referents are construed as 'anaphoric' in the sense that they need to be bound to already available discourse referents. Now these antecedents can be provided by *other* attitudes. Thus I may believe that I will catch *a fish* and decide to have *it* fried for dinner. Here the free discourse referent in the content of the decision is anaphoric on the discourse referent introduced by the indefinite in the content of the belief. So a discourse referent introduced by an attitude is parasitically exploited by another attitude (Maier 2015).

'Anchors' are a special type of attitude, which introduces discourse referents *on the basis of acquaintance relations*. Qua attitudes, they have a label (ANCH) and a content (represented by a DRS). For example, a perceptual anchor introduces a discourse referent and ascribes it a set of conditions corresponding to the perceptual relation itself (the fact that the object is currently seen) *and* the information which can be immediately gained through the perceptual relation (that the object is red, that it is a mailbox). The discourse referent thus introduced by the anchor can serve as the antecedent of 'free' discourse referents in the DRS which represents the content of an attitude. Thus the subject sees a red mailbox and judges that *it* may contain a letter.

What difference is there between the mental file framework as expounded in Recanati (2012) and DRT ? Maier says it is the informational content of the file. Anchors only contain descriptive conditions pertaining to the perceptual relation and information immediately gained from it. In contrast, mental files also contain everything the subject believes regarding the reference of the file (and similarly for the other attitudes). The content of beliefs, for Maier as for Pryor and Losada, is distributed over the entire belief system— the set of attitudes with the 'BEL' label and an associated DRS. Because attitudes can be parasitic on other attitudes (including anchors), the content of belief is actually distributed over the entire

*attitude* system. But the informational content of *anchors* is restricted to reference-fixing information for the file. This is a form of Descriptivism, as Maier acknowledges : the relation that fixes reference is reflected in the descriptive content of the file.

### 2.5 Mental files and mindreading in psychological development

Perner and Leahy emphasize the role which *sortals* play in file thinking. (See Carey 2009 on the transition from spatio-temporal individuation to sortal individuation of objects in psychological development.) ‘Individuation requires individuation-as’, they say. Since one cannot refer to an object in thought or speech without individuating it, objects that are represented in thought or speech are always represented-as something. Perner and Leahy say that each file comes with a label which provides a *conceptual perspective* on the reference of the file.

Perner and Leahy assume that the same file cannot bear distinct labels. Each file bears a single label : its header. So if one and the same object is conceptualized in different ways, under alternative sortals, *different files will be opened for that object* – corresponding to each of the ‘guises’ under which the object may appear (see Crimmins 1998 for a related view). The files, however, will be *linked*, enabling free information flow between them (Recanati 2012). For example, there will be different files for Paderewski the pianist and Paderewski the politician in a knowledgeable adult’s mind, and they will be linked to each other so that information in each file is accessible from the other. Linking is an important operation precisely because objects are associated with distinct guises and it seems appropriate to use the mental file construct to represent the guises, as Perner and Leahy do. (Many theorists use mental files to deal with Clark Kent/Superman, and Perner’s and Leahy’s example of the rubber die which is also an eraser is similar.)

Perner and Leahy show that children before the age of 4 do not master linking and therefore fail to represent identities. ‘Children younger than about four years can anchor each file to the same object but cannot represent this fact, i.e. cannot represent the identity of the object referred to by the files.’ They open as many files as there are perspectives on an object, but do not link the files. They treat different guises of the object (e.g. the rubber die which is also an eraser) as distinct objects, associated with distinct bodies of information which do not communicate with each other. This, Perner and Leahy argue, explains why three year olds fail both the false belief task and the alternative naming task: both tasks crucially require linking.

(For tasks involving the representation of belief, a special category of files is needed : so-called ‘vicarious files’ or ‘indexed files’. Perner and Leahy argue that the child represents the attitudes of others with respect to objects by means of vicarious files about these objects, indexed to these people. Before passing the false belief task, however, they fail to link these vicarious files to their regular files about the object.)

The last paper in the issue is Agnes Kovacs’s ‘Belief Files in Theory of Mind Reasoning’. Kovacs is concerned with ‘two-system’ approaches to the development of mindreading. To reconcile the robust findings about failures on the false-belief task by three year olds and the more recent findings about the existence of ‘early mindreading’, many theorists have argued for a distinction between implicit (early) and explicit (late) forms of mindreading. Kovacs agrees that such a distinction is needed, but she resists the suggestion that early mindreading might not involve the representation of belief as such, but only of relations between the agent and the object (or between the agent, the object and the location in the standard false belief task). She argues that spontaneous tracking and monitoring of people’s beliefs starts very early and operates on *belief representations*, just like mature Theory of Mind does. And she offers hypotheses regarding the format of these belief representations :

The precondition for spontaneous mindreading is a basic representational structure, called here the ‘belief file’, that enables implicit ToM processes to store information about other agents’ beliefs in a format supporting efficient encoding and updating.

Belief files as Kovacs conceives of them are not files about objects but, rather, about situations (e.g. where the ball is), so the content of belief files is propositional rather than predicative. A belief file is indexed to a particular agent (the believer) but the belief file involves separate components for the believer and the content of the belief ascribed to him or her : having a separate component makes it possible to update only the belief content (while keeping the agent fixed). Kovacs argues that belief files can be ‘empty’, by which she means that the content of the ascribed belief may not be fully specified but may involve placeholders, making it possible to update the belief file at a later stage. The possibility of empty belief files shows that ‘young infants can sustain attributed belief representations independently from their first person representations about reality’, a crucial point in the debate over early mindreading.

## References

- Ariel, M. (1988) Referring and Accessibility. *Journal of linguistics* 24: 65-87.
- Bach, K. (1987) *Thought and Reference*. Oxford: Clarendon Press.
- Baillargeon, R. (1995) Physical Reasoning in Infancy. In M. Gazzaniga (ed.), *The Cognitive Neurosciences*, 181-204. Cambridge, Mass.: MIT Press.
- Breckenridge, W. and Magidor, O. (2012) Arbitrary Reference. *Philosophical Studies* 158: 377-400.
- Burke, L. (1952) On the Tunnel Effect. *Quarterly Journal of Experimental Psychology* 4: 121-138.
- Campbell, J. (1987) Is Sense Transparent? *Proceedings of the Aristotelian Society* 88: 273-92.
- Carey, S. (2009) *The Origin of Concepts*. New York: Oxford University Press.
- Carey, S., & Xu, F. (2001) Infants' Knowledge of Objects: Beyond Object Files and Object Tracking. *Cognition* 80 : 179-213.
- Chafe, W. (1976) Givenness, Contrastiveness, Definiteness, Subjects, Topics, and Points of View. In C. Li (ed.) *Subject and Topic*, 27-55. London: Academic Press.
- Chafe, W. (1994) *Discourse, Consciousness, and Time: The Flow and Displacement of Conscious Experience in Speaking and Writing*. Chicago: University of Chicago Press.
- Crimmins, M. (1998) Hesperus and Phosphorus: Sense, Pretence, and Reference. *Philosophical Review* 107: 1-47.
- Devitt, M. (2014) Lest Auld Acquaintance Be Forgotten. *Mind & Language* 29: 475-484.
- Erteschik-Shir, N. (1997) *The Dynamics of Focus Structure*. Cambridge: Cambridge University Press.
- Fine, K. (1983) A Defence of Arbitrary Objects. *Proceedings of the Aristotelian Society* 57: 55-77.

- Fine, K. (2007) *Semantic Relationism*. Oxford: Blackwell.
- Flombaum, J., Scholl, B., & Santos, L. (2009) Spatiotemporal Priority as a Fundamental Principle of Object Persistence. In B. M. Hood & L. Santos (eds.) *The Origins of Object Knowledge*, 135-164. Oxford: Oxford University Press.
- Fodor, J. (1983) *The Modularity of Mind*. Cambridge, Mass: MIT Press/Bradford Books.
- Fodor, J. & Pylyshyn, Z. (2014) *Minds without meanings: An essay on the content of concepts*. Cambridge, Mass.: MIT Press.
- Grice, P. (1969) Vacuous Names. In D. Davidson and J. Hintikka (eds.) *Words and Objections*, 118-45. Dordrecht: Reidel.
- Gundel, J. K., Hedberg, N., & Zacharski, R. (1993) Cognitive Status and the Form of Referring Expressions in Discourse. *Language* 69: 274-307.
- Kahneman, D., Treisman, A., & Gibbs, B. J. (1992) The Reviewing of Object Files: Object-specific Integration of Information. *Cognitive psychology* 24: 175-219.
- Kamp, H. (1981) A Theory of Truth and Semantic Representation. In P. Portner and B. H. Partee (eds.), *Formal Semantics: the Essential Readings*, 189-222. Oxford: Blackwell.
- Kamp, H. (1990) Prolegomena to a Structural Theory of Belief and Other Attitudes. In C. Anderson and J. Owens (eds.) *Propositional attitudes*, 27-90. Stanford: CSLI Lecture Notes.
- Kamp, H. (2015) Using Proper Names as Intermediaries between Labelled Entity Representations. *Erkenntnis* 80: 263-312.
- Kaplan, D. (1989) Demonstratives. In J. Almog et al. (eds.) *Themes from Kaplan*, 481-563. New York: Oxford University Press.
- Karttunen, L. (1976) Discourse Referents. *Syntax and Semantics* 7: 363-85.
- Lambrecht, K. (1996) *Information Structure and Sentence Form: Topic, Focus, and the Mental Representations of Discourse Referents*. Cambridge: Cambridge University Press.

- Landman, F. (1986) Pegs and Alecs. In J. Halpern (ed.) *Theoretical Aspects of Reasoning about Knowledge*, 45-61. Los Altos: Morgan Kaufmann.
- Maier, E. (2015) Parasitic Attitudes. *Linguistics and Philosophy* 38: 205-36.
- Millikan, R. (2012) Are There Mental Indexicals and Demonstratives? *Philosophical Perspectives* 26: 217-34.
- Noles, N., Scholl, B., & Mitroff, S. (2005) The Persistence of Object File Representations. *Perception & Psychophysics* 67: 324-334.
- Perry, J. (1979) The Problem of the Essential Indexical. *Noûs* 13: 3-21.
- Perry, J. (1980) A Problem about Continued Belief. *Pacific Philosophical Quarterly* 61: 317-32.
- Pinillos, A. (2011) Coreference and Meaning. *Philosophical Studies* 154: 301-24.
- Prince, E. (1981) Toward a Taxonomy of Given-New Information. In P.Cole (ed.) *Radical Pragmatics*, 223-255. New York: Academic Press.
- Pylyshyn, Z. (2000) Situating Vision in the World. *Trends in Cognitive Science* 4 : 197-207.
- Pylyshyn, Z. (2001) Visual Indexes, Preconceptual Objects, and Situated Vision. *Cognition* 80: 127-158.
- Pylyshyn, Z. (2007) *Things and Places: How the Mind Connects with the World*. Cambridge, MASS: MIT Press/Bradford Books.
- Recanati, F. (1993) *Direct Reference: From Language to Thought*. Oxford: Blackwell.
- Recanati, F. (2004) *Literal Meaning*. Cambridge: Cambridge University Press.
- Recanati, F. (2012) *Mental Files*. Oxford: Oxford University Press.
- Recanati, F. (2013) Mental Files: Replies to My Critics. *Disputatio* 36: 427-43.
- Recanati, F. (2016) *Mental Files in Flux*. Oxford: Oxford University Press.



- Reinhart, T. (1981) Pragmatics and Linguistics: An Analysis of Sentence Topics. *Philosophica* 27: 53-94.
- Rips, L., & Hespos, S. (2015) Divisions of the Physical World: Concepts of Objects and Substances. [\*Psychological Bulletin\*](#) 141:786-811.
- Scholl, B. J., & Leslie, A. M. (1999) Explaining the Infant's Object Concept: Beyond the Perception/Cognition Dichotomy. In E. Lepore & Z. Pylyshyn (eds.) *What is cognitive science*, 26-73. Oxford: Blackwell.
- Schiffer, S. (1978) The Basis of Reference. *Erkenntnis* 13: 171-206.
- Scholl, B. (2001) Objects and Attention: the State of the Art. *Cognition* 80: 1-46.
- Spelke, E. (1990) Principles of Object Perception. *Cognitive science* 14: 29-56.
- Sperber, D. and Wilson, D. (1998) The Mapping between the Mental and the Public Lexicon. In P. Carruthers and J. Boucher (eds.) *Language and Thought: Interdisciplinary Themes*, 184-200. Cambridge: Cambridge University Press.
- Strawson, P. (1971) *Logico-Linguistic Papers*. London: Methuen.
- Strawson, P. (1974) *Subject and Predicate in Logic and Grammar*. London: Methuen.
- Vallduví, E. (1992) *The Informational Component*. Garland: New York.
- Williamson, T. (2007) *The Philosophy of Philosophy*. Oxford: Blackwell.
- Woodfield, A. (1991) Conceptions. *Mind* 100: 547-572.