The Metaphysics of Mental Files

Simon Prosser

University of St Andrews

ABSTRACT

There is much to be said for a diachronic or interpersonal individuation of singular modes of presentation (MOPs) in terms of a criterion of epistemic transparency between thought tokens. This way of individuating MOPs has been discussed recently within the *mental files* framework, though the issues discussed here arise for all theories that individuate MOPs in terms of relations among tokens. All such theories face objections concerning apparent failures of the transitivity of the 'same MOP' relation. For mental files, these transitivity failures most obviously occur because mental files can merge or undergo fission. In this paper I argue that this problem is easily resolved once mental files are properly construed as *continuants*, whose metaphysics is analogous to that of persons or physical objects. All continuants can undergo fission or fusion, leading to similar transitivity problems, but there are well-established theories of persistence that accommodate this. I suggest that, in particular, the stage theory best suits the purposes of a continuant theory of MOPs.

INTRODUCTION

The mental files framework provides a way of modelling non-descriptive singular modes of presentation (MOPs). It first appears in the work of Paul Grice (1969), Michael Lockwood (1971) and Peter Strawson (1974), was developed in more detail by John Perry (1980 and elsewhere), and has more recently received a substantial further defence and elaboration by François Recanati (1993, 2012, 2016).[1] The general idea is that the mind contains functional analogues of 'files' or 'dossiers' that compartmentalise information about specific items. If, for example, S believes that Hesperus is bright and that Phosphorus is pretty, this is because S's mind has a file associated with the

---

[1] See the preface to Recanati 2012 for a more detailed account of the history of mental files.

name 'Hesperus' that contains the predicate 'is bright' and a file associated with the name 'Phosphorus' that contains the predicate 'is pretty'. Even if, in fact, Hesperus is Phosphorus, the predicates remain in their respective files unless S comes to believe that Hesperus is Phosphorus. This compartmentalisation of information explains why a subject can hold conflicting attitudes toward the same state of affairs, or can attribute incompatible properties to the same object, without irrationality. One way to think of this is that, from S's point of view, it could have turned out that the files for Hesperus and Phosphorus concerned different objects, for there are possible worlds containing an intrinsic duplicate of S in which this would be true; and S would only be irrational if S's overall world view was inconsistent in relation to every way the world might have turned out to be. Mental files thus constitute singular MOPs; the mere numerical difference between the files constitutes the difference in MOPs. The information that they contain in no way constitutes the identity of the MOP.

For all its advantages, however, the mental files framework faces an objection concerning the way it individuates MOPs. This concerns failures of transitivity of the 'same MOP' relation that arise when mental files merge of undergo fission. In this paper I shall argue that this objection is easily dealt with once we have a proper understanding of the metaphysics of mental files. Mental files are best thought of as *continuants*, just like persons or physical objects. All continuants can undergo fission or fusion, and there are well-known theories in metaphysics that accommodate this. I shall argue that the stage theory, in particular, is best suited to the purposes of modelling MOPs. With the metaphysics properly understood, it will then become apparent that the mental files framework and the relationalist view (Fine 2007, Heck 2012) need not be seen as rivals.

The mental files framework is an example of a family of views that accept that there are MOPs but do not individuate them in terms of content.[2] Like the

---

[2] Other examples include Kaplan 1990, Fodor 1998, Schroeter 2012, Schroeter and Schroeter 2014, 2016, Sainsbury and Tye 2012. Note that some of these philosophers regard the mental files framework as a way of modelling their view despite not presenting their views chiefly in that form. Fine 2007 and Heck 2012 also fit into the broad family of views that I have in mind, and I shall argue below that their relationist views are in fact quite close to the mental files view, once the latter is combined with the correct metaphysics of persistence.

mental files view, all such theories construe sameness of MOP in terms of certain relations that hold between tokens. Because such relations can be one-many, all such theories are subject to just the same objections concerning fission and fusion, and the solution that I propose for mental files will apply equally to all of them. I would like to emphasise, then, that although I shall largely discuss matters in terms of mental files, my aim is to defend a broader notion of what constitutes a MOP.

THE INDIVIDUATION OF MENTAL FILES

Frege (1956, 1966) provides a clear criterion for differences in MOPs in a single subject at a single time. If, at time t, a rational subject S can believe that a is F while denying that b is F, or withholding judgment about it, then 'a' and 'b' correspond to different MOPs for S at t. Where S cannot take differing attitudes, it is usually assumed that the tokens of 'a' and 'b' correspond to the same MOP. But this criterion does not tell us what we should say about the interpersonal or diachronic cases where the question is what it takes for two different subjects, or the same subject at different times, to think of an object, o, under the same or different MOP.

The mental file theory, along with other views in the same family, usually deals with diachronic identity in terms of a relation of epistemic transparency between tokens. This can be understood in terms of John Campbell's (1987, 1994) notion of *trading on identity*. Consider the following inference, in which the subscripts are used only to designate tokens:

a_1 is F
a_2 is G

Something is both F and G

One can often infer the conclusion without needing a further premise identifying the references of the two tokens of 'a'. That is, rather than adding the identity of the references of 'a_1' and 'a_2' as an additional assumption, one 'trades' on their co-reference in making the inference. As Campbell argues, it is hard to see how a rational thinker could get by without trading on identity; and, in any case, there seems no doubt that we do as a matter of fact reason in this way extremely often. Campbell argues that it is possible to trade on identity in this way just when the two tokens of 'a' express the same MOP – for this is exactly the kind of rational relation among thoughts that MOPs are supposed to capture. Indeed it is precisely the fact that we do not trade on identity between tokens of 'Hesperus' and tokens of 'Phosphorus' that blocks the inference from 'Hesperus is bright' to 'Phosphorus is bright' in the absence of the assumption that Hesperus is Phosphorus. By definition, then, two tokens are of the same file, and hence the same MOP, just if it is possible for the subject to trade on identity between them.[3] For brevity, I shall say that two tokens are *transparently related* just when it is possible for the rational subject to trade on identity between them; so two tokens are of the same MOP if and only if they are transparently related.

The mental files framework seems a natural partner for an individuation of MOPs in terms of transparency. In the above inference, for example, the fact that the subject can trade on identity is connected with the fact that the tokens 'a_1' and 'a_2' are both associated with the same mental file. Because the predicates 'is F' and 'is G' both appear in the same file they can be conjoined and predicated without further inference. If one of the tokens were associated with a different a different file then the inference would require at least one additional premise in order to associate a predicate in one file with a predicate in another file. Note, however, that since 'same mental file' is itself to be understood in terms of the transparency relation, the notion of a file does not in itself carry much explanatory weight. Rather, at least on my understanding,

---

[3] Recanati (2012, 2016) also uses trading on identity as his criterion for identity of MOPs; and others with views of the same general kind, while not always making explicit reference to Campbell's notion, normally describe things in a corresponding way (see e.g. Heck 2012: 155, Schroeter 2012: 180).

the notion of a file should just be thought of as a metaphor that helps capture the way in which information is organised in the subject's mind. The fact that the information is organised in that way must be given a separate explanation. In particular, the fact that two predicates appear in the same file is to be explained in terms of that file standing in a certain kind of ongoing epistemic relation to the reference. I shall say more about this below.

There is far more to be said about mental files, for example concerning what we should say about cases of reference failure, or cases in which there is transparency from the subject's point of view yet an unnoticed change of reference (for example in a case where a perceived object is switched for a duplicate without the subject noticing). I shall not, however, go further into the details of this, since the problem with which I shall be concerned, and its solution, will be the same for all plausible variants of the view.[4]

THE IDENTITY OF MOPS OVER TIME AND BETWEEN PEOPLE

Provided we restrict our attention to synchronic intrapersonal inferences and Frege cases, the mental files framework, understood as described above, works well. But what about the individuation of MOPs over time, or between different thinkers? As mentioned above, most advocates of the mental files framework hold that transparency is the correct criterion in dynamic cases as well, though there are some slight complications to discuss (see below). Several – myself included – also hold that transparency is the right criterion in interpersonal cases.[5]

First, then, the diachronic case. As Gareth Evans (1982: 235) pointed out, it is a requirement for being a rational thinker that one's thoughts at one time have a rational bearing on one's thoughts at subsequent times. Were this not the case, there would be little point in retaining information. We can capture this

---

[4] For further discussion see Recanati 2016, especially concerning the difference between the weak and strong versions of coreference *de jure*. See also Prosser forthcoming.

[5] For the interpersonal case see for example Dickie and Rattan 2010, Prosser forthcoming, Recanati 2016, Schroeter 2012.

rational bearing in terms of transparency. Suppose that, at t_1, one thinks that Hesperus is bright and that Phosphorus is pretty. Given that Hesperus is in fact Phosphorus, but assuming that one does not know this, why is it that at t_2, in the absence of new information, one thinks that Hesperus is bright and Phosphorus is pretty instead of thinking, say, that Hesperus is pretty and Phosphorus is bright? It is because the different tokens of 'Hesperus' are associated with the same mental file, and thus the same MOP, as are the different tokens of 'Phosphorus'. Each predicate remains in its file, and one thus retains the very same beliefs, individuated at the level of MOPs. We can think of this in terms of transparency because retention of a belief over time is like a limiting case of inference: one infers 'Hesperus$_2$ is bright' from 'Hesperus$_1$ is bright' without the need for further premises, thus trading on the identity of reference of the tokens 'Hesperus$_1$' and 'Hesperus$_2$'.

One could, of course, try to resist this view. One could insist that MOPs occurring at different times are necessarily distinct. But this is a difficult position to sustain. It would require an explanation of why retained predicates 'stick' to one singular term over time (e.g. 'Hesperus') rather than another. But why try to defend that view, when the assumption that MOPs are retained through time provides such a straightforward and intuitive account? Unless there is a very strong objection to the retention of files, and thus MOPs, over time, the view should be accepted. It provides a way to capture the rational structure of the thinker's thoughts over time. I take it that correctly modelling the thinker's rational thought processes is the main reason for individuating MOPs in the first place.

Traditionally, MOPs were thought to be individuated in terms of their epistemic properties. What I mean by 'epistemic properties' is this: two token MOPs differ with respect to their epistemic properties if, and only if, they differ with respect to the epistemic conditions under which the subject would be warranted in making a judgment that deploys the MOP. Hence, according to traditional theories of this kind, two token thoughts were of different types, and thus involved different MOPs, if and only if they differed with respect to the conditions under which the subject would be warranted in making the

relevant judgment about the reference. This view became popular because of an excessive focus on the synchronic intrapersonal Frege cases. If a rational subject, S, at a time *t*, can believe that *a* is *F* while denying that *b* is *F*, this can only be because S is unaware that *a* = *b*. But it can appear possible to S that *a* and *b* are different objects only if S forms judgments about *a* and *b* under different epistemic circumstances. It must, for example, appear to S that *a* is to the left while *b* is to the right, or that *a* is the morning star while *b* is the evening star. If this were not the case – if the conditions under which S would predicate 'is *F*' of *a* and *b* were exactly the same – then it would never come about that S believed something about *a* without also believing the same about *b*.[6]

What this shows is that if S makes different judgments concerning '*a*' and '*b*' at a given time then the MOPs associated with '*a*' and '*b*' differ epistemically *at that time*. But all too often a far stronger conclusion has been drawn, to the effect that if, at time *t*, for subject S, '*a*' and '*b*' have different epistemic properties E_1 and E_2 respectively, then the MOPs associated with '*a*' and '*b*' are *essentially* characterised by E_1 and E_2. Descriptivist theories of singular MOPs are perhaps the best-known example, but the same applies to theories that see MOPs in terms of other epistemic notions such as recognitional capacities (and it also applies to theories that individuate MOPs in terms of specific kinds of epistemic relations, on which see below).

The epistemic properties of a mental file can change over time, however, without there being any reason to think that there is a change in the identity of the file, provided there is transparency throughout.[7] Consider an analogy with physical objects. Assuming Leibniz's principle of the identity of indiscernibles,

---

[6] Would it be possible to think of two objects as distinct even if one believed exactly the same about both of them, and even if both MOPs had the same epistemic properties? Perhaps the mere distinctness of two mental files, albeit with exactly the same contents, would make this possible. It depends on whether Leibniz's principle of the identity of indiscernibles is implicitly hard-wired into one's psychology. If so, then when there is no epistemic difference between two files, the files should be considered as having merged into a single one (and the subject would thereby take it that there was only one object). I take it to be an empirical question whether our brains are wired up in this way; a question whose answer is likely to be independent of whether the principle of the identity of indiscernibles is true, but which might shed some light on our intuitions regarding the principle.

[7] Laura Schroeter (2012) makes an analogous point for the interpersonal case in terms of what she calls the *Flexibility Constraint* – that different speakers may refer to the same object and keep track of this in conversation despite large differences in what they believe about the object.

if two physical objects are distinct at time t then they differ in at least one property (which can include spatial location). Suppose that, at t, objects o_1 and o_2 differ because they instantiate different properties P_1 and P_2 respectively. It clearly does not follow that o_1 and o_2 are *essentially* characterised by their possession of P_1 and P_2. Many of the properties of a physical object can change over time, and could even change such that at some later time o_1 instantiated P_2 while o_2 instantiated P_1. It is an essential fact about physical objects that they possess a certain range of properties (location, size, shape etc.) at any given time, and it is essential that two distinct objects differ with respect to at least one of those properties at any given time, but it is not essential to the identity of a given physical object that it possess a specific set of properties. What makes an object the same one from one time to another is that its states depend in certain systematic ways on its earlier states, not that it continues to possess the same set of properties.[8] If transparency is the criterion for sameness of mental file then we should think of the relation between a mental file and its epistemic properties as analogous to the relation between a physical object and its physical properties. It is essential to a mental file that at any given time it has certain epistemic properties that determine the conditions under which the subject is warranted in making judgments the deploy the mental file, but no mental file can be individuated just in terms of such properties, as they may change over time. In some cases transparency may require a temporary overlap in epistemic properties, such as when an object is identified as the evening star, then as both the evening star and the morning star, and subsequently only as the morning star (perhaps because the first description has been forgotten). But this need not always be true. For example one may track a visually perceived object as it moves around, gradually changes colour, shape, and so on, such that there is no overlap in its perceived properties from one time to another, yet one's thoughts about it may be diachronically transparent.

---

[8] I have glossed over some issues relating to sortal properties. Perhaps being a table lamp is essential to table lamps; perhaps no table lamp can become an egg. But the point made above will still apply to a wide range of other properties of table lamps, and no table lamp can be individuated solely by instantiating the property of being a table lamp.

I am using the phrase 'epistemic property' as an umbrella term, covering any kind of property that determines the epistemic circumstances under which it is appropriate for the subject to deploy the MOP. I reject any theory according to which MOPs have their epistemic properties essentially, though, as explained above, I do accept MOPs have specific epistemic properties for a given subject at a given time. Many mental file theorists hold that the reference of a mental file is determined by a specific kind of epistemic property: a certain kind of epistemic relation in which the file stands to its reference. Recanati (2012) calls these 'Epistemically Rewarding' or 'ER' relations. There is room for different accounts of these relations while remaining within the same broad framework, but the general idea is that each file will be 'tuned' to a specific entity – the reference of the file – by virtue of the epistemic relation between the entity and the mental file. Information acquired by virtue of that relation, and placed in the file, will normally be information about that entity. Recanati distinguishes several types of ER relations, and types files according to the ER relations that they exploit. For example a demonstrative file is one that normally contains information about an object that the subject is perceptually tracking, a HERE file is one that contains information about the current location of the subject, a NOW file is one that contains information about the time that is cotemporaneous with the file, and so on.

Recanati (2015, 2016) agrees with Papineau (2006, 2013), Ball (2014), and Ninan (2014) that for many purposes we need to talk about files that persist through changes in their ER relations. He calls these persisting files 'encyclopedia entries'; they stand in higher-order ER relations to their references, consisting in the possession of one of the first-order ER relations. But he holds that we also need a finer-grained individuation wherein files are typed by their ER relations. An encyclopedia entry exists when a series of fine-grained files with different ER relations occurs, such that each one is 'converted' into the next, to which it is transparently related. For example, a demonstrative file, based on a current perception of an object, may be converted into a different kind of file when the object is no longer perceived, provided the new file is associated with a new ER relation, as might occur when the subject

acquires a recognitional capacity for the object. An encyclopedia entry consists in the series of fine-grained files of this kind.

If there is genuine theoretical utility in the finer-grained individuation of files then I have no objection to positing them, alongside the coarser-grained encyclopedia entries that are individuated by epistemic transparency relations as described above. But I am yet to be convinced of their utility. Recanati sometimes seems to assume that fine-grained files are needed in order to deal with Frege cases (see for example the exchange between Recanati (2015) and Onofri (2015)). But this is not correct, much as it would not be correct to hold that physical objects must be individuated in a fine-grained way in order to account for the fact that there can be two distinct physical objects at the same time. Rather than positing a series of distinct physical objects whenever their properties change it is much simpler to speak of a single object whose properties change over time; and, similarly, much simpler to speak of a single persisting file whose epistemic properties change over time. If we accept that there are coarse-grained files then we must already accept that they have epistemic properties at any given time, and that these epistemic properties change over time. I cannot see what is added by also claiming that there are distinct files, of the fine-grained type, at different times.

Another putative reason for appealing to fine-grained files would be to account for systematic differences in behaviour associated with files based on different ER relations. This comes out most clearly with indexicals, though similar points can be made about other cases. When one thinks of a place as *here* at t_1 one may behave differently than if one subsequently thinks of the same place as *there* at t_2, even if one keeps track of the place throughout, such that one's *here* and *there* thoughts are related transparently. One might, for example, think 'it is always raining here' and, as a result, one opens one's umbrella. But one may then move away from the place, keeping track of it throughout. Barring loss of memory or new information, one will be disposed to hold, of the same place, that 'it is always raining there'. The latter thought will be transparently related to one's earlier thought, but will not dispose one to open one's umbrella. Perry (1977, 1979) has given several well-known examples that

illustrate such systematic differences in behaviour associated with different indexicals.

A difference in rational behaviour must be explained in terms of a difference in mental states. It may therefore seem tempting to associate this difference in thoughts with the difference in ER relations associated with *here* and *there* (one judges that it is *F here* when it is *F* in one's own location, whereas one judges that it is *F there* when it is *F* at some other location). Many philosophers have thus concluded that *here* and *there* must always correspond to different fine-grained MOPs of a place. Perhaps the unstructured linguistic form of indexical pronouns such as 'here', 'now' and 'I' has contributed to a view that each different indexical is associated with a different unstructured thought component with a special cognitive role. In the mental files framework this would be understood in terms of types of files that have special cognitive roles.

In my view, this common line of thought has been a significant wrong turn. Given a common set of desires, systematic differences in behaviour can always be explained in terms of the way the subject takes the world to be. For any indexical pronoun, IND, which refers to *o*, if the subject expresses a belief using words of the form 'IND is *F*' then the subject believes of *o*, that it is *F*. But S also believes something else, as manifested by the choice of indexical term. If S uses the word 'now' in referring to time *t*, for example, then S believes that *t* is present. Similarly, when S uses the word 'here' in referring to location *l*, S believes that *l* is *hereabouts*. This is a relation between S and *l*: roughly speaking, the relation such that S is located within *l*. Finally, when S uses 'I' in referring to S, S believes that S stands in a certain relation to the person S thinks about. Arguably (though the details won't matter here) this relation involves S being the person about whom S receives information in certain ways (e.g. through proprioception), and whose behaviour can be controlled in various ways by S. Let us call predicates such as 'is present' and 'is hereabouts' *egocentric predicates*. Egocentric predicates typically have a monadic form: one can think in terms of *past*, *present*, *future*, *hereabouts*, *left*, *right* and so on without needing to articulate

to oneself the fact that these are really relations to oneself.[9] In linguistic terms, one can think of egocentric predicates – or, at least, their fully articulated equivalents – as presuppositions of an utterance containing an indexical pronoun. A use of 'here' presupposes that the speaker is located at the place that is the reference of the pronoun, for example.

Consequently there is a difference in what S believes when S goes from thinking 'it is always raining here' to thinking 'it is always raining there', about the same place. In both cases, S believes, of location *l*, that it is always raining at *l*. Since, if S has kept track of *l*, S's thought tokens will be transparently related, there seems no reason to deny that there is a single persisting mental file whose reference is *l*. But, in addition, S goes from believing that *l* is hereabouts to believing that *l* is thereabouts (where 'thereabouts' refers to some other relation between S and *l*, different from the 'hereabouts' relation). This is clearly a difference in how S takes the world to be, and thus a difference in what S believes. It is *this* difference that explains the difference in S's behaviour; not a difference in the MOP under which S thinks of *l*. And since this is all part of what S believes about *l*, and all under the same MOP of *l*, the predicates 'is hereabouts' and 'is thereabouts' are among the contents of the persisting mental file for *l*. Finer-grained files, and the associated finer-grained MOPs and thoughts, thus have no role to play in the story.[10, 11]

Note that I am not claiming that S has beliefs concerning the ER relations of mental files in every case. Where an indexical pronoun is used, it presupposes, or connotes, the relation specified by an egocentric predicate.[12] There need be no similar presuppositions associated with other uses of mental files. But there

---

[9] I give a more detailed account of the relations in question, and the sense in which the mental representation of them is monadic, in Prosser 2015, where I also argue for their essential role in action.

[10] I have defended the above claims about the cognitive dynamics of thoughts expressed using indexicals in more detail in Prosser 2005. See also Prosser forthcoming for the equivalent claims about communication and shared indexical thoughts.

[11] In recent work, Recanati (2016) seems willing to appeal to the presuppositions of the use of an indexical when discussing the corresponding problem for shared indexical thoughts. Given that this seems similar to my own solution in both the dynamic and interpersonal cases (on which see below and Prosser forthcoming), it is not clear to me why Recanati should not accept the same move in the dynamic case.

[12] However even here there are cases in which the egocentric predicate fails to correctly apply to the object. See Prosser 2012 for details.

are also no systematic behavioural differences associated with such cases. Suppose that S has two different names for the same person, Smith, and, at a given time, the two corresponding files are associated with two different ways of visually recognising Smith. For example, perhaps one file is activated when S sees Smith from one side, whereas the other file is activated when S sees Smith from the other side (we can suppose that S has never seen both sides of Smith within a single perceptual episode, and therefore never realised that there is just one person). S might be quite unaware of how she goes about recognising Smith in the two cases, and therefore although the two files differ in their ER relations to Smith, this does not entail a difference in the contents of the files, which would constitute a difference in belief. The two files are not associated with systematic differences in behaviour in the way that different indexical pronouns are. Any differences in behaviour are likely to result from contingent differences in the contents of the files (i.e. differences in what S believes about Smith under the different MOPs).

Consequently I can see no need to posit fine-grained files, individuated by their ER relations. It is important to stress, however, that if finer-grained files have a theoretical utility that I have missed, this will not affect the arguments that follow, provided there is still a role for persisting, coarse-grained files, or encyclopedia entries. It is the latter notion of a mental file that individuates MOPs purely in terms of relations among tokens, and for which the transitivity problems most clearly arise (though insofar as similar transitivity problems can arise on Recanati's finer-grained individuation of MOPs, the stage-theoretic solution proposed below should still apply).

As I shall explain below, the diachronic intrapersonal case is sufficient to illustrate the transitivity problem with which we shall be concerned. Since interpersonal cases provide particularly clear examples of the problem, however, I shall say a little about the interpersonal individuation of MOPs. If MOPs are to be individuated by mental files then for two people to think of an object under the same MOP is for their thought tokens to involve tokens of the same mental file. But what does it mean for mental files realised in separate brains to be tokens of the 'same' file? The obvious way to approach the problem

is to keep firmly in mind that an individuation of MOPs should capture the kinds of epistemic relations that exist between token thoughts, as described above. Where token MOPs are related transparently they are tokens of the same MOP, and thus the token mental files to which they correspond should be typed together.[13]

Here is one suggestion for how this transparency might occur.[14] For there to be epistemic transparency between the token thoughts of two people, it must be possible for one person's belief to lead to the other person predicating the same property of the same object by trading on the identity of reference between the token singular terms of the two people. This happens in those cases in which, when a speaker makes an utterance, there is no need for the audience to entertain an interpretive hypothesis about the reference of the singular term. Thus, for example, when the speaker says 'Hesperus is bright', someone who hears this may come to believe what they would express using the words 'Hesperus is bright', without ever needing to employ an explicit assumption that the other speaker's token of 'Hesperus' referred the same object as their own (or to Hesperus, as they would put it). Instead, coreference is presupposed. This happens precisely when there is a shared language. It is the interpersonal analogue of the intrapersonal phenomenon of trading on identity described above.

The precise details are not important for present purposes, however. The view just described, for both the intrapersonal diachronic case and for the synchronic interpersonal case, says that what makes it the case that two tokens are of the same MOP is that a certain kind of relation holds between the tokens. This contrasts with theories according to which two tokens are of the same MOP if and only if both tokens have an epistemic property in common (a property other than that of being related to other tokens in a certain way, such as the property of being applicable when a certain descriptive condition is met).

---

[13] Dickie and Rattan (2010) in effect argue for a criterion of transparency for interpersonal sameness of MOPs, albeit they do not relate this explicitly to the mental files approach.

[14] The following view is described in more detail in Prosser forthcoming, and is endorsed by Recanati (2016). A similar view has been developed by Laura Schroeter and François Schroeter (Schroeter 2012; Schroeter and Schroeter 2014).

The problem described below will arise for any view that belongs to the first, relational category; and the solution will be of the same kind in each case.

TRANSITIVITY PROBLEMS

So far, so good. But upon closer inspection there is a prima facie problem with an individuation of MOPs based on epistemic transparency, or other similar relations between tokens. Several authors have pointed out alleged counterexamples involving a failure of transitivity of the 'same file' relation or, equivalently, the 'same MOP' relation (Soames 1989, 1994, Fine 2007, Pinillos 2011). In some such cases it can be shown that transitivity does not in fact fail.[15] I shall focus on cases involving fission and fusion, however, as these provide the clearest cases in which transitivity does fail.

I take it for granted that the identity relation is transitive. So if three token thoughts, A, B and C, are such that A involves the same MOP as B, and B involves the same MOP as C, then A should involve the same MOP as C. The counterexamples are cases where this fails.

Case 1. Communication problems. Consider three persons, S_1, S_2 and S_3. S_1 expresses a singular thought by saying 'Smith is tall', in the presence of S_2. S_2 recognises the name 'Smith', and is thus able to trade on identity. S_2 thus comes to entertain a thought expressible as 'Smith is tall' without ever needing to consider whether S_1 and S_2's tokens of 'Smith' refer to the same person. By the above arguments, then, S_1 comes to entertain the same fine-grained thought, containing the same singular MOP, as S_2. The same happens when S_1 speaks to S_3, so by the same reasoning S_1 and S_3's 'Smith' MOPs are also the same. But when S_2 and S_3 try to communicate using the name 'Smith', something goes wrong. Perhaps the three speakers' pronunciation of the name varies in just such a way that although S_2 and S_3 can both recognise S_1's utterance as

---

[15] See Recanati (2016: chapter 3) for details.

containing the name 'Smith', for which they each have a pre-existing file, neither can recognise the other's pronunciation. Perhaps whenever S_2 says 'Smith', S_3 hears it as 'Snith', a name for which S_3 has no pre-existing file. In order to make sense of S_2's utterance, S_3 needs an interpretive hypothesis, such as '"Snith" (as uttered by S_2) stands for Smith'. Since S_3 thus fails to trade on the identity of reference between S_2 and S_3's tokens of 'Smith', S_2 and S_3 do not associate the same MOP with the name. But this violates the transitivity of the 'same MOP' relation.

Case 2. Fission of files through fission of persons. Suppose that a person undergoes perfectly symmetrical fission, and the 'parent' is psychologically continuous with both of the resulting 'offspring'. As is well-known, such cases create transitivity problems for psychological continuity theories of personal identity. If the parent is psychologically continuous with both offspring then, according to the psychological continuity theory, the parent should be identical with both offspring, yet they are not psychologically continuous with one another, and therefore not identical with one another. Something very similar can happen with the mental files of those undergoing fission. The files of the offspring may be transparently related to those of the parent; yet it does not follow that the files of both offspring will be transparently related to one another. They may fail to recognise one another's pronunciation of a name, for example, as described in case 1. In fact one reason for mentioning case 2 in addition to case 1 is to illustrate the close relation between transparency theories of mental files and psychological continuity theories of personal identity.

Cases 3 and 4. Intrapersonal fusion of files. Even in the intrapersonal case similar problems arise. For example, case 3: When someone learns that Hesperus is Phosphorus their two files may eventually merge into one. The resulting file will inherit predicates transparently from each of the original files, yet the original files were distinct from one another. Note that in such cases there should normally be no contradiction between the contents of the two original files; for if the subject had believed that Hesperus was *F* while believing that

Phosphorus was not *F* then it would not have been rational to identify Hesperus with Phosphorus. Even if files containing contradictory predicates could somehow merge, however, I see no objection to saying that the fused file would transparently inherit their contents, and the subject would then be left with contradictory beliefs that must somehow be resolved.

For reasons related to this latter complication, it is less clear whether it is possible for a file to undergo the relevant kind of intrapersonal fission. Suppose, for example, that you came to believe, mistakenly, that what you had thought of as London was in fact two different cities. One was the capital city of the UK; the other was the most populous city in the UK, and so on. You might thus separate the predicates from your original parent file into two distinct offspring files, depending on which city you took each predicate to apply to. Since there was in fact only one city, all three files would have the same reference. However, since it would not be the case that all predicates from the parent file were automatically transferred to both of the offspring files, there would not be diachronic transparency, and hence no clear case for holding that the subject retained a MOP.

But perhaps there is a way to construct a more elaborate case. Suppose (case 4) that a person undergoes symmetrical fission as described above, such that every parent file has two offspring files, one in each offspring person. Suppose the offspring persons then fuse again. Let the original file be F1, and the two offspring files be F2 and F3. We can imagine that during the period of separation an epistemic difference arises between F2 and F3, and that as a result, when the persons fuse, the files do not. Instead, the fused person has two distinct files, F4 and F5. In that case there would be transparency along the chains F1-F2-F4 and F1-F3-F5, but no transparency between F4 and F5, and so once again a failure of transitivity.

Perhaps many further cases exist. Anyway, I take it that there are sufficient cases to show that there is a genuine problem with transitivity failure when we take transparency to be sufficient for sameness of MOP in co-referential files.

TRANSPARENCY AND PERSONAL IDENTITY

The account of MOPs outlined above is clearly analogous to a common strand of thought about personal identity. In fact, arguably, it is more than an analogy (though nothing in the rest of the paper depends on the connections that I shall make in this section being correct). While there is still much debate over the correct criterion for the identity of persons over time, psychological continuity theories remain popular and do seem to capture something important about us. According to psychological continuity theories, person 1, who exists at time t_1, is identical with person 2, who exists at a later time t_2 if and only if there is psychological continuity between the thoughts of person 1 and person 2 (there are of course well-known problems with such theories, which I shall come to below). There are different ideas about the details of what psychological continuity amounts to. But I take the general idea to be that the continuation of a person over time is the continuation of a single rational mind. Person 1's psychological state must influence person 2's psychological state, and in the right way. It is not enough that person 1's mental states merely cause person 2's mental states; that could happen in cases where persons 1 and 2 were different people. The natural way to understand what counts as the 'right' way is that person 2's thoughts are transparently related to person 1's thoughts. Memory is transparent in this way, for example. If, at t_1, one thinks 'the Grand Canyon is in Arizona', remembering this at t_2 does not require one to explicitly identify the reference of one's current 'Grand Canyon' token with that of one's earlier 'Grand Canyon' token. One doesn't think about the earlier token at all; one simply remembers that the Grand Canyon is in Arizona. One thus trades on the identity of reference between the two tokens of 'Grand Canyon', and one's earlier and later thoughts are transparently related. So, given this understanding of psychological continuity, a psychological continuity theory of personal identity entails that it is a necessary condition for the persistence of a person over time that there is at least some degree of transparency between the earlier and later files in the mind of the persisting person. I shall not attempt to state the sufficient conditions. Following Parfit (1971), I acknowledge that

both psychological continuity and personal survival may be a matter of degree, and this might be at least partially understood in terms of the proportion of files that are transparently related between different stages of the person.[16]

There may be a further analogy in the synchronic case. Suppose there were perfect transparency between the mental files realised in two different brains. This would require that the contents of both sets of files were perfectly synchronised. For any file A_1 in brain 1, there would be a corresponding file, A_2 in brain 2, such that if any predicate appeared in one of the files it would immediately appear in the other, and would do so without the mediation of an interpretive premise. In that case, all of the thoughts based in the two brains would remain perfectly synchronised. Suppose that there were two brains thus synchronised, and with the same desires as well (which would perhaps be necessary if the brains were to remain synchronised). In such a case it would become plausible to say that there was really only one person present, albeit a person with two perfectly synchronised brains. Two minds might thus merge into one. By analogy with survival over time, we might think of merging as a matter of degree. Consequently, given that one typical aim of communication is to synchronise files, and given that it can take place transparently as described above, we can think of transparent communication as quite literally a partial merging of minds, albeit usually on quite a limited scale.[17]

A STAGE THEORY OF MENTAL FILES

It is well known that the possibilities of personal fission or fusion create problems for psychological continuity theories of personal identity (Shoemaker 1963, Wiggins 1967, Parfit 1971). If the post-fission offspring are both psychologically continuous with the pre-fission parent then the psychological continuity theory entails that both offspring are numerically identical with the

---

[16] I have glossed over Parfit's (1971) distinction between *continuity* and *connectedness* here. Strictly speaking, it is what Parfit calls 'connectedness' that would require transparency.

[17] The ideas discussed in this paragraph are obviously somewhat underdeveloped; I include them only as food for thought.

parent, yet one cannot be identical with two. The problem can be thought of in terms of transitivity; identity is transitive, yet since the parent is psychologically continuous with both offspring, while the offspring are not psychologically continuous with each other, the relation of psychological continuity fails to be transitive. The same kind of problem is encountered with almost any kind of continuant: continuity relations are typically one-many, and therefore cannot be extensionally equivalent to the identity relation in all possible scenarios.

The root of the problem is the assumption that the persistence of a continuant consists in the numerical identity of an entity that exists in its entirety at one time with an entity that exists in its entirety at another time. To persist in this way is to *endure*.[18] But although the endurance theory of personal identity still has advocates, many philosophers have sought to deal with fission, fusion and other problem cases through a different account of persistence, according to which a continuant persists by having numerically distinct stages at different times.

This comes in two different forms. The *perdurance* theory says that a continuant is a temporally extended entity that has different parts at different times. Fission and fusion cases are dealt with by holding that distinct continuants can share temporal parts, in a way that is analogous to the sharing of spatial parts by spatially extended objects, such as apartments that share a wall. Consider, for example, a case of fission in which a temporal part P, at time t_1, is succeeded after fission by temporal parts Q and R at time t_2, each of which is related to P in the way in which successive temporal parts of a single continuant are normally related. The standard perdurance account says that there are two persisting entities: O_1, consisting of P and Q, and O_2, consisting of P and R (see Lewis 1976). This has the consequence that what exists at t_1 is a part of two different continuants, and hence there is a sense in which there are

---

[18] See Merricks 1999 for a further argument, not concerning transitivity, to the effect that the numerical identity of persons over time cannot be constituted by the psychological continuity relation.

two different continuants present at t_1. This may seem counterintuitive, especially when the continuants in question are persons.

Lewis (1976: 64) responds to this objection by arguing that the way we count objects in such situations is non-standard: in effect, we count how many objects there are that are non-identical at that time. This claim is not uncontroversial, but I shall not attempt to settle it here (for dissent from Lewis see Sider 1996). Instead I shall move directly to consider the second theory, which is known as the *stage* theory (Hawley 2001, Sider 1996, 2001). According to the stage theory, what we call a continuant is in fact a series of distinct entities called 'stages', each of which is identical with one of the temporal parts posited by perdurance theory. Some metaphysicians regard the stage theory and the perdurance theory as terminological variants of the same metaphysical view. Insofar as there is a genuine metaphysical difference it concerns the metaphysics of the part-whole relation. If some sets of entities compose to form a whole, whereas others do not, then the theories may be seen as making genuinely different claims. If, on the other hand, every set of entities composes to form a whole (as mereological universalism claims) then the stage theory is a terminological variant of the perdurance theory; and if no set of entities ever composes a whole (as mereological nihilism claims) then the stage theory is strictly the correct theory, though again the perdurance theory might be interpreted as a terminological variant.

The stage theory is analogous to a counterpart theory of modality (Lewis 1986). A person, or other persisting entity, is a momentary stage. When I say that I was once a child, the truth expressed by this is that there was an earlier person-stage, to whom I stand in the appropriate *I-relation* (the relation in which one temporal part of a person stands to another, according to the perdurance theory), and that person-stage has (or had) the property of being a child. The I-relation can be thought of as loosely analogous to the modal counterpart relation. Different types of I-relation apply to different types of entity, according to their persistence conditions. For example when a clay statue undergoes a process of fission via a gradual replacement of clay parts with bronze parts, this results in a bronze statue and a lump of clay that is not

a statue. In such a case, the original object stands in the *clay-I-relation* to the subsequent lump of clay and also stands in the *statue-I-relation* to the subsequent bronze statue. Where there is fission of an entity into more than one entity of the same type – a person becomes two persons, for example – then the parent entity stands in the same I-relation to both of the offspring. This has the advantage of avoiding the consequence of the perdurance theory that the parent is a part of two different temporally-extended entities, and thus avoids the consequence that two different entities (such as two different persons) are present even before fission takes place.

Given that mental files are continuants, it is very straightforward to model their persistence according to either the perdurance theory or the stage theory. On either account, what we call a persisting mental file comprises a series of numerically distinct momentary or very short-lived files, which I shall call *file-stages*, each of which stands in the transparency relation to at least one later file-stage. We can think of file-stages as what were called file-tokens in the discussion above. The transparency relation is thus the I-relation for mental files; I shall say that transparently related file-stages are *T-related*, for short. If the stages are instantaneous, and if time is continuous, then it would not be strictly correct to say that each file-stage is T-related to the *next* file-stage, since there is no next file-stage in a continuum. But we can still say that for any two file-stages occurring at different times, both involve the same persisting file, and thus the same MOP, if and only if the two file-stages are T-related. Moreover, although the interpersonal case is not diachronic, and thus provides no strict analogy with the metaphysics of persistence, there seems no obvious reason not to apply the same general principles. Hence we can say that two people are thinking of an object under the same MOP at the same time if and only if the two file-stages are T-related.

When a mental file undergoes fission the parent file, F1, is T-related to both of the offspring files F2 and F3, but F2 and F3 may or may not be T-related to each other. But there is no resulting problem; the transparency relation is intransitive and one-many, but since the different file-stages are numerically distinct from one another there can be no objection to the theory based on a

failure of transitivity. This will be true whether we regard the post-fission files as parts of two different temporally-extended files, F1-F2 and F1-F3, as the perdurance theory has it, or else as two files-stages F2 and F3, as the stage theory has it.

Given their close relationship I shall not argue strongly for one or other of the perdurance theory or the stage theory. But I do feel that the stage theory has one significant advantage for modelling MOPs. As noted above, the perdurance theory has the prima facie drawback that it entails a sense in which there are two different entities present even before fission, because the pre-fission parent is a part of two different temporally-extended composites. This seems particularly undesirable for a theory of MOPs; for, prior to fission, the subject cannot rationally take two different attitudes to the same state of affairs, as we should normally expect if the subject entertains two different MOPs. The stage theory avoids this consequence, as only one file-stage, and thus only one MOP, is present prior to fission. There are probably ways for the perdurance theory to be modified in order to get around this difficulty; it is probably not a decisive objection to the perdurance theory (for example see again Lewis's (1976: 64) discussion of a non-standard way to count entities in such cases). But given the straightforward simplicity of the stage theory I take it to be at least the prima facie more satisfactory way to model the diachronic and interpersonal metaphysics of MOPs.

Adopting the file-stage theory does not amount to abandoning talk of the same MOP. The stage theory is not revisionary of most ordinary talk; it does not recommend that we stop saying that this article was written by a single person, for example. There is something that we capture when we talk in this way, which would be lost if we were simply to say instead that the article was written by a series of different people. The same applies to MOPs. But the theory does require that in certain contexts, when speaking carefully, we understand talk of a single persisting entity as shorthand for talk of a series of I-related stages. My proposal can be thought of as at least loosely analogous to Derek Parfit's (1971) proposal about personal identity. The lesson that Parfit drew from the fission and fusion cases is that survival, which he construes in

terms of psychological connectedness, is not identity. Identity is a one-one relation; whereas survival can be one-many. The stage theory provides a good way to model this, as Sider (1996) observes. Similarly, the mental files theorist should accept that the 'survival' of a file, or of a MOP, is not identity. But there are still many contexts in which it is helpful to talk about the 'same' person, and similarly many contexts in which it is helpful to talk about the 'same' MOP.[19, 20]

The stage theory helps with a further problem that arises for the mental files theory. Consider again the interpersonal case. I have suggested that MOPs are shared just when there is the ability to trade on identity; person A says 'Hesperus is bright', and person B comes to believe what B would express using the words 'Hesperus is bright', without ever having to entertain thoughts about the interpretation of A's token of 'Hesperus'. I noted above that the transitivity of the 'same MOP' relation fails in cases where two speakers fail to recognise each other's tokens of a word as being of the same word, and thus fail to trade

---

[19] A different approach to transitivity problems is given by Laura Schroeter and François Schroeter (2014: 16-18). In addition to the appearance of de jure coreference, Schroeter and Schroeter add a condition of 'congruence' on the individuation of MOPs. This deals with cases in which there is a difference in reference between tokens despite epistemic transparency between them, leading to failures of transitivity. The congruence condition amounts to saying that tokens are only of the same concept when they belong to the same *representational tradition*, the set of tokens gathered together as the unit for semantic interpretation by the best overall interpretation. However this does not address the transitivity problems described above, which do not involve any shift in reference. Perhaps one could save transitivity by appeal to a finer-grained individuation of representational traditions, such that, for example, there is only sameness of MOP in non-branching cases. But the equivalent move for persons and physical objects (i.e. persistence is non-branching continuity) seems worryingly ad hoc, and the theory would lead to a proliferation of MOPs that would not clearly provide the most helpful way to model the rational relations among thought tokens. For my own view on how to deal with cases of reference shift see Prosser forthcoming.

[20] R. M. Sainsbury and Michael Tye (2012) have developed a theory of concepts that they call *originalism*, according to which concepts are abstract continuants individuated by their origins, such that 'necessarily, concept C1 = concept C2 iff the originating use of C1 = the originating use of C2' (2012: 44). This puts their view within the broad family of theories that I wish to defend, but I disagree with certain aspects of their specific version. In particular, when discussing fission and fusion cases (2012: 66-68) they take it for granted that two new concepts are created when there is fission. Although this allows them to treat the 'same MOP' relation as an identity relation it means that in such cases diachronic transparency cannot be accounted for in terms of a retained MOP. Moreover, since the offspring concepts are distinct from one another, interpersonal transparency also cannot be accounted for in terms of a shared MOP. Given that the reason for individuating MOPs in the first place was to capture rational relations among token thoughts, this seems to me to sacrifice too much. Replacing their implicitly endurantist metaphysics of concepts with a stage theory, as described above, solves the problem.

on identity, even though each can trade on identity with the same third speaker. But there must also be the possibility of asymmetric cases, in which A can recognise B's token of 'Hesperus', and thus trades on identity when forming beliefs based on B's testimony, yet B does not recognise A's token, and thus fails to trade on identity. It seems that we should have to say that A shared B's MOP, yet B does not share A's MOP, which clearly conflicts with the symmetry of the identity relation.

While such cases are most easily noticed in the interpersonal case, they can also occur intrapersonally. Suppose there were a person who was disposed to infer b is F directly from a is F, without an identifying premise such as '$a = b$', but who was not disposed to infer a is F from b is F without the further premise. Clearly there would be something wrong with the rationality of such a person; but there is no obvious reason why it should not be possible for such a person to exist. Should we say that this person's 'a' tokens express the same MOP as their 'b' tokens? There are four possible answers: (1) sameness of MOP requires symmetry, so the MOPs are not shared. But then how do we account for the subject's trading on identity, albeit in just one direction? (2) The MOPs are shared. But then what stops the subject trading on identity in both directions? (3) There is no fact of the matter; when someone is irrational to such a degree, the individuation of their thoughts is indeterminate. (4) We should abandon the assumption that the 'same MOP' relation really involves numerical identity, and instead accept that what we call the 'same MOP' relation is an asymmetric relation. In the interpersonal case, A can think what B is thinking, even though B does not think what A is thinking.

Given the problems posed by (1) and (2), only answers (3) and (4) seem to have much merit. Either of (3) or (4) is compatible with the overall view that I am advocating. But I lean toward (4), especially for the interpersonal case. If we are going to accept that the 'same MOP' relation is not really a relation of identity, why not accept that it can be asymmetrical? It is worth noting that this issue has not generally arisen in discussion of the stage theory. The stage theory was developed in order to deal with the metaphysics of persistence. The I-relations of the stage theory normally hold between earlier and later stages, and

involve causation. Given that causal relations are directional, the question of whether the I-relation holds in both directions doesn't really arise – it holds from earlier to later, and does not need to hold from later to earlier. It is asymmetric by default.

But the individuation of MOPs provides us with something new: a case in which the 'sameness' I-relation – specifically the T-relation – can hold synchronically, and in either or both of two directions. In such a case, why not interpret the stage theory in such a way as to allow asymmetric synchronic I-relations? This allows us to capture the full range of epistemic relations among token thoughts; whereas (3), according to which all we can say about asymmetric cases is there is no answer as to whether one token involves the same MOP as the other, fails to tell us about the transparency that exists in one direction, and therefore leaves something out.

Note that adopting the file-stage theory does not mean adopting fine-grained files in Recanati's sense. Were we to individuate files by ER relations then we would need to consider both the transparency relation and 'same ER relation' relations between file-stages. But if we can explain everything in terms of transparency relations, the 'same ER relation' relation is superfluous.

Thomas Hofweber and J. David Velleman (2011) have recently suggested that we can make a new distinction between two different ways of persisting, without needing to make reference to temporal parts or stages, and they suggest that the words 'endurance' and 'perdurance' would be better used to capture this distinction. While I do not entirely agree with their complaints about the standard endurance/perdurance distinction, I do think the distinction they make is useful, and has particular relevance to the case at hand. Their proposal, at least roughly, is that for perduring entities the identity of the entity at a given time depends on relations to entities existing at other times, whereas the identity of an enduring entity is constituted entirely by what exists at a single time. So, for example, standard psychological continuity accounts of personal identity, such as Locke's, count as perdurance theories because they hold that the identity of a person, P, at a time depends the identity of the earlier person to whom P stands in the relation of psychological continuity. On the

other hand, a theory according to which the identity of a person depends solely on the identity of a soul, whose identity is intrinsic to it, would count as an endurance theory. This maps nicely on to the distinction, made above, between traditional theories that identify a MOP by the epistemic properties of the token, and those, like the mental files theory, that appeal to relations between tokens. The mental files theory, and other such 'relational' theories, entail that MOPs perdure, in Hofweber and Velleman's sense. A descriptivist theory, on the other hand, according to which the MOP contained in a token thought depends entirely on the descriptive content of the token MOP, would count as an endurance theory.[21]

RELATIONISM

I shall finish with a brief comment on the relation between the file-stage theory and relationist theories such as *semantic relationism* (Fine 2007) or *formal relationism* (Heck 2012). I shall not attempt to adjudicate between these two versions of relationism, as my purpose is only to point out that relationism of some kind can almost certainly be regarded as compatible with some version of the file-stage theory.[22] For the purposes of illustration I shall focus on Fine's theory, though similar comments should apply to Heck's theory (and in some respects Heck's theory may be more straightforwardly compatible with the file-stage theory, given that the relations in question are not at the semantic level on Heck's view). According to Fine, Frege puzzles and similar puzzles are generated by a false view that he calls *intrinsicalism*. This is the view that semantic differences can only be due to differences in semantic features intrinsic to individual token thoughts or utterances. For example, the difference between the obvious *de jure* coreference of the names in 'Cicero is Cicero' and the merely *de facto* coreference of the names in 'Cicero is Tully' must be

---

[21] Schroeter and Schroeter (2016) also make a similar distinction between what they call the 'binding' and 'matching' models of concepts (MOPs), though they do not connect this with Hofweber and Velleman's distinction.

[22] On the difference between semantic and formal relationism, however, see Heck 2014.

explained in terms of semantic differences intrinsic to the tokens of 'Cicero' and 'Tully'. By analogy (2007: 2) Fine cites the fact that two identical twins look alike because they share intrinsic features. By contrast, Fine recommends a view according to which such matters are to be explained in terms of semantic relations among tokens that are not reducible to the intrinsic semantic features of the tokens. The main such relation is what Fine calls *coordination*. Fine puts this in terms of representing the reference as the same, as opposed to representing the reference as *being* the same, which would involve having the thought that they are the same (as, for example, when one thinks 'Cicero is Tully'). Importantly, the semantic relations with which Fine is concerned need not be transitive.

Fine says relatively little about mental files, though he raises an objection concerning the possibility of files 'merging' (undergoing fusion). The objection seems to be largely that 'it is not that the merged file represents the individual as the same as the earlier files, since that would require that the earlier files represent the individual as the same' (2007: 69). As distinct files, they do not. As I read it, this seems to be precisely the kind of transitivity objection to which the above arguments are addressed.

There are similarities between Fine's distinction between intrinsicalism and relationism, and Hofweber and Velleman's distinction between endurance and perdurance. Given this, and the role of T-relations in the file-stage theory, we can construe relationist views as terminological variants of the file-stage theory. Relationist theories deal in individual utterance tokens and the relations between them, and the stage theory does much the same. But given the stage theory, we can see that talk of the *same* file, or, more broadly, the *same* MOP (for those who reject mental files but accept a view in the same family), remains precisely as acceptable as talk of the same person. Such talk need not be rejected, provided it is understood as a kind of shorthand for the stage-theory metaphysics.[23]

REFERENCES

Ball, D. 2015. Indexicality, Transparency, and Mental Files. *Inquiry* 58: 353-367.

Campbell, J. 1987. Is Sense Transparent? *Proceedings of the Aristotelian Society*, 88: 273-292.

Campbell, J. 1994. *Past, Space and Self*. Cambridge, MA: MIT Press.

Dickie, I, and Rattan, G. 2010. Sense, Communication, and Rational Engagement.l*Dialectica*, 64: 131-151.

Evans, G. 1982. *The Varieties of Reference*. Edited by John McDowell. Oxford: Clarendon Press.

Fodor, J. A. 1998. *Concepts: Where Cognitive Science Went Wrong*. Oxford: Clarendon Press.

Fine, K. 2007. *Semantic Relationism*. Oxford: Blackwell.

Frege, G. 1956. The Thought: A Logical Inquiry. Translated by A. M. and Marcelle Quinton. *Mind*, 65: 289-311.

Frege, G. 1966. On Sense and Reference. In *Translations from the Philosophical Writings of Gottlob Frege*. Translated by P. Geach and M. Black. 3rd edn. Oxford: Blackwell: 56-78. (Translation of Frege, G. 1892. Über Sinn und Bedeutung. *Zeitschrift für Philosophie und Philosophische Kritik*, 100: 25-50).

Grice, H. P. 1969. Vacuous Names. In D. Davidson and J. Hintikka (eds.) *Words and Objections*. Dordrecht: Reidel: 118-45.

Hawley, K. 2001. *How Things Persist*. Oxford: Oxford University Press.

Heck, R. 2012. Solving Frege's Puzzle. *Journal of Philosophy*, 109: 132-174.

Heck, R. 2014. In Defense of Formal Relationism. *Thought*, 3: 243-250.

Hofweber, T., and Velleman, J. D. 2011. How to Endure. *The Philosophical Quarterly*, 61: 37-57.

Kaplan, D. 1990. Words. *Aristotelian Society Supplementary Volume*, 64: 93-119.

Lewis, David 1976. Survival and Identity. In Amélie O. Rorty (ed.), *The Identities of Persons*. Berkeley: University of California Press: 17-40. Reprinted in David Lewis, *Philosophical Papers, Volume 1*. Oxford: Oxford University Press: 55-77.

Lewis, David. 1986. *On the Plurality of Worlds*. Oxford: Blackwell.

Lockwood, M. 1971. Identity and Reference. In M. K. Munitz (ed.) *Identity and Individuation*. New York: New York University Press: 199-211.

Merricks, T. 1999. Endurance, Psychological Continuity, and the Importance of Personal Identity. *Philosophy and Phenomenological Research*, 59: 983-997.

Ninan, D. 2015. On Recanti's *Mental Files*. *Inquiry* 58: 368-377.

Onofri, A. 2015. Mental Files and Rational Inferences. *Inquiry* 58: 378-392.

Papineau, D. 2006. Phenomenal and Perceptual Concepts. In T. Alter and S. Walter (eds.) *Phenomenal Concepts and Phenomenal Knowledge: New Essays on Consciousness and Physicalism*. New York: Oxford University Press: 111-144.

Papineau, D. 2013. Comments on François Recanati's *Mental Files*: Doubts about Indexicality. In F. Salis (ed.) *Book Symposium on François Recanati's* Mental Files. *Disputatio*, 5: 159-76.

Parfit, D. 1971. Personal Identity. *Philosophical Review*, 80: 3-27.

Perry, J. 1977. Frege on Demonstratives. *The Philosophical Review*, 86: 474-497.

Perry, John 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: 301-324.

Prosser, S. 2005. Cognitive Dynamics and Indexicals. *Mind & Language*, 20: 369-391.

Prosser, S. 2012. Sources of Immunity to Error Through Misidentification. In Prosser, S. and Recanati, F. (eds.) *Immunity to Error Through Misidentification: New Essays*. Cambridge: Cambridge University Press: 158-179.

Prosser, S. 2015. Why are Indexicals Essential? *Proceedings of the Aristotelian Society*, 115: 211-233.

Prosser, S. forthcoming. Shared Modes of Presentation. *Mind & Language*.

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

Recanati, F. 2012. *Mental Files*. Oxford: Oxford University Press.

Recanati, F. 2015. Replies. *Inquiry* 58: 408-437.

Recanati, F. 2016. *Mental Files in Flux*. Oxford: Oxford University Press.

Sainsbury, R. M, and Tye, M. 2012. *Seven Puzzles of Thought*. Oxford: Oxford

University Press.

Shoemaker, S. 1963. *Self-Knowledge and Self-Identity* Ithaca, NY: Cornell University Press.

Sider, T. 1996. All the World's a Stage. *Australasian Journal of Philosophy*, 74: 433-453.

Schroeter, L, 2012. Bootstrapping Our Way to Samesaying. *Synthese*, 189: 177-197.

Schroeter, L, and Schroeter, F. 2014. Normative Concepts: A Connectedness Model. *Philosophers' Imprint*, 14: 1-26.

Schroeter, L, and Schroeter, F. 2016. Semantic Deference Versus Semantic Coordination. *American Philosophical Quarterly*, 53: 193-210.

Sider, T. 2001. *Four Dimensionalism*. Oxford: Oxford University Press.

Soames, S. 1989. Presupposition. In D Gabbay and F. Guenther (eds.) *Handbook of Philosophical Logic, Vol. IV*. Dordrecht: Reidel: 553-616.

Soames, S. 1994. Attitudes and Anaphora. *Philosophical Perspectives*, 8: 251-72.

Strawson, P. 1974. *Subject and Predicate in Logic and Grammar*. London: Methuen. Reprinted by Ashgate Press, 2004.

Wiggins, D. 1967. *Identity and Spatio-Temporal Continuity*. Oxford: Blackwell.