The Mental Files Theory of Singular Thought: A Psychological Perspective

Michael Murez, Joulia Smortchkova, Brent Strickland

Michael Murez, Joulia Smortchkova, Brent Strickland

The Mental Files Theory of Singular Thought: A Psychological Perspective

DRAFT VERSION submitted on August 1st 2015 for volume on *Mental Files and Singular Thought.* Contracted with Oxford University Press, edited by J. Genone, R.Goodman & N. Kroll. (Please do not cite without permission). Contact: michael.murez@ens.fr

Abstract: We argue that the most ambitious version of the mental files theory of singular thought, according to which mental files are a wide-ranging psychological natural kind underlying all and only singular thinking, is unsupported by the available psychological data. Nevertheless, critical examination of the theory from a psychological perspective opens up promising avenues for research, especially concerning the relationship between our perceptual capacity to individuate and track basic individuals, and our higher level capacities for singular thought.

1. Introducing singular thought: the philosophical explanandum

Many thoughts are *general*: they are about properties or categories. For example, the *generic* thought that wombats are cute is paradigmatically general.[1] So are *quantificational* thoughts, like the thoughts that all (or some) wombats are cute. However, we also have *singular* thoughts, thoughts about particular individuals. For example, if one sees an unfamiliar creature and judges that it is cute, whatever *it* is, one's thought is singular: One is thinking about *that* very individual, which one sees but cannot recognize. Such perception-based *demonstrative* thoughts are the most basic variety of singular thoughts. But other varieties exist, which are based on memory or communication rather than perception. For example, until now, you had never heard of Patrick, the "wombassador" from Ballarat. Yet now we have told you about Patrick the wombat, you can think about *him* – for instance, you may wonder whether Patrick is cute.

There are also *descriptive* thoughts, thoughts about whatever satisfies a certain description: for example, the thought that the world's oldest wombat (whichever one that is) is cute. Since the world's oldest wombat happens to be Patrick, this thought is also, in a sense, about a particular individual: what makes the thought accurate is *Patrick's* being cute. Yet unlike genuine singular thoughts, descriptive thoughts are not about individuals *per se*. They refer to individuals only indirectly, by representing attributes they possess (e.g., wombathood and elderliness). If some other wombat had turned out the oldest, the same thought would have been about *it* instead. Hence, descriptive thoughts about individuals are considered a kind of general (quantificational) thought: they are really about small categories, with at most one member, rather than directly about the particular category members themselves.

[1] A generic thought is about a category or kind. While classifying generic thoughts as general is common, for example in encyclopedia entries (Crawford, 2013) they are sometimes described as singular, for the following reasons: a) occasionally the term 'singular' is used, misleadingly, for all nondescriptive thoughts b) sometimes kinds are identified with special sorts of individuals which occupy multiple disconnected locations at the same time. We find this use of 'singular' misleading, given the straightforward fact that multiple individuals can belong to the same kind. But in any case, classifying generic thoughts as singular would only strengthen the arguments we make in this chapter.

The distinction between descriptive and nondescriptive[2] 'modes of aboutness' (Sellars, 1949) applies not only to thoughts about individuals, but also to thoughts about general entities like properties ('the color of daffodils' vs. 'yellow') or kinds ('mammals whose scat is cubical' vs. 'wombats'). So, underlying the dichotomy of singular and general thought, we have a finer-grained fourfold classification of thoughts along two dimensions – the nature of their subject matter or referential content, and whether or not their reference is determined through satisfaction of a descriptive condition:

	Nondescriptive	Descriptive
Individual	Singular thought	General thought
Property or kind	General thought	General thought

Among these varieties of thought, *singular* thought has been claimed to be special in various ways, e.g., with respect to its epistemology, metaphysics, or role in the naturalization of content. However, from an opposing perspective, singular thought has also been claimed by so-called *descriptivists* (Lewis, 1984; Schiffer, 1978; Searle, 1983) to bear no essential difference to descriptive – hence general – thought.

One of the main descriptivist arguments stems from *Frege Cases*, in which subjects rationally attribute contradictory properties to the same referent. For example, someone who does not realize that Fat Pat is Patrick can rationally judge that Patrick is cute, yet concurrently deny that Fat Pat is.

To account for such cases, philosophers have proposed that thoughts can involve distinct "modes of presentation" (MOPs) of the same referents, which are functionally defined by *Frege's Constraint*: One can rationally think, of some entity, that it has a certain property and its negation just in case one represents the entity under different MOPs (Schiffer, 1990). MOPs also satisfy a *Transparency Constraint*: one can know purely introspectively whether the MOPs under which one is thinking of something are identical or different (Boghossian, 1994).

Descriptivists claim only descriptions can fill the role of MOPs. Since MOPs feature in cases of apparently singular thinking, descriptivists conclude that even seemingly singular thoughts actually go through descriptions. The contrast between singular and descriptive thought dissolves.

However, there are well-known independent reasons to reject descriptivism (Kripke, 1980). Among these, it seems one could be wrong, ignorant, or change opinions about how to describe an individual, yet still think about *it*. For example, one can think about Patrick without being able to tell him qualitatively apart from a homonymous twin,[3] or keep track of

[2] This distinction is also known following Bach (1987) as the 'satisfactional/nonsatisfactional' distinction.
[3] Even infants countenance the possibility of indistinguishable individuals being different, for example by

him over time as one's conception of him evolves: "It's a bird, it's a plane, it's Superman – no, it's Patrick the wombat!" (Kahneman, Treisman, & Gibbs, 1992).

Consequently, rather than reduce singular to descriptive thought, many philosophers have proposed to distinguish two varieties of MOP, descriptive and *nondescriptive* (NMOPs). This move promises to address Frege Cases while avoiding descriptivism: Singular thoughts are thoughts in which individuals are represented under NMOPs, rather than descriptions. But what *are* NMOPs?

2. Mental file theory: a possible explanans from philosophy, in need of empirical support

The mental file theory of singular thought (MFT) claims that NMOPs reduce to mental files, which psychologically explain the capacity to entertain singular thoughts. On this view, singular thinking *just is* file-thinking: that the representational vehicles of singular thoughts are files is what fundamentally differentiates them from general or descriptive ones.[4]

'Mental files' are mental representations, whose functions are to refer to some entity, and also to collect, store, and render re-accessible information about the entity. A file is structurally complex, being composed of the file *itself*, and a collection of descriptive *entries* 'inside' the file, which represent attributes co-predicated of its referent. For example, one's mental file about Patrick might 'contain' entries like "is a wombat", "is from Ballarat", etc.

Defenders of MFT conceive of files as *independent* from the descriptive entries they contain, in several respects:

1) What *determines the identity* of a file, i.e., makes some representation(s) numerically the same/different mental file(s), is not the identity of its entries.

2) What *determines the reference* of a file, i.e., files refer through causal, contextual, or historical relations, such as what Recanati (2012) calls 'epistemically rewarding relations' (ERs) as opposed to referring via descriptions.

3) What *determines the entries* inside a file, i.e., makes it the case that a certain descriptive representation is associated with a certain file, is that, according to Recanati (2012), information gets filed together because it is gained through a common ER.

4) What *determines the access* to a file, i.e., makes it the case that one retrieves, activates, or 'deploys' a certain file, is not activation of a certain entry inside the file. Files are not addressed *via* their descriptive content. Again, Recanati (2012) claims that ERs rather than

---

preferring original possessions over perfect duplicates (Hood & Bloom, 2008).

[4] There are many different versions of MFT, as well as many closely related views (Bach, 1987, 2010; Crane, 2011; García-Carpintero, 2000; Jeshion, 2010; Lawlor, 2001; Montague, 2011; Perry, 1993; Recanati, 1993, 2010; Sainsbury, 2005; Sawyer, 2012; Taylor, 2003). The version we focus on in this paper is distinctive in taking file-thinking to be constitutively tied to singular thought: Jeshion (2010, p. 132): "Thinking about an individual from a mental file is constitutive of singular thinking about that individual". Jeshion (2010, p. 129): "One thinks a singular thought by thinking through or via a mental file that one has about the particular object. By contrast, descriptive thoughts occur discretely in cognition, disconnected from any mental file." Recanati (2012, p. 13): "the singularist distinction [between singular and descriptive thought] reduces to the distinction between two kinds of sense or mode of presentation, descriptive and non-descriptive". Recanati (2012, p. 34): "A non-descriptive mode of presentation, I claim, is nothing but a mental file."

entries play the relevant role: A file's function is to be deployed so long as the subject stands in the appropriate ER to its referent.

According to MFT, files' independence from their descriptive entries is what makes them especially suited to the NMOP-role. Differences between coreferential singular thoughts are explained by numerical differences between coreferential files themselves, rather than associated descriptions (entries in files). For example, I can have distinct files/NMOPs for Patrick and his qualitatively indistinguishable twin, even if no information inside distinguishes them. Since files so conceived can refer to entities which do not satisfy their entries, they also address the problems of ignorance, error, and change plaguing descriptivism. For example, I can have a file which (persistently) refers to Patrick because its entries causally derive from him, even though they depict him as a groundhog named 'Patricia', or despite the fact that they change as my opinions about him evolve.

This attractive picture faces a major obstacle. If files are psychologically *real*, then merely sketching an account in which they fit the task-description of essentially singular representations *a priori* is insufficient. What tells us that such files not only exist, but also behave the way MFT says, rather than like *general* files, which refer to properties of kinds (Fodor, 2008; Prinz, 2005; Schiffer, 1996), or like *descriptive* files, whose reference, entries, and access are determined by a description (Goodman, forthcoming)? It will not do to respond that by 'mental file', MFT just *means* file-like representations that are appropriately singular. The issue is not how we use the technical expression 'mental file', but whether an empirically well-motivated notion in its vicinity can explain or support, rather than merely *label*, the distinction between singular mental representations and general or descriptive ones.

These objections suggest singular thinking cannot be reduced to file-thinking from the armchair. However, this is not a knockdown objection, if MFT is understood as an *empirical* thesis, according to which representations of a certain *psychological natural kind*, traditionally referred to as 'mental files', are *a posteriori* constitutive of our capacity for singular thought. Conceptual analysis may not reveal all the "central features of file-hood" (Goodman, forthcoming) so understood: some of these features, potentially including *singularity*, may not be deducible from files' functional-theoretical role. That is one of the characteristics of psychological natural kinds.

2.2 Psychological natural kindhood

Thoughts not only have contents, which they are about, but also *vehicles*, which we (our brains) think them with: symbolic representations, with 'formal' properties in virtue of which they play causal-functional roles in the computational processes that constitute thinking.

Each token singular thought has some vehicle, which collectively can be grouped under the label 'singular representations'. From a semantic perspective, these representations serve a common semantically-specified function or task: they are for representing individuals nondescriptively. But from a formal perspective, is there anything theoretically interesting that unites this class of 'singular representations', and distinguishes them from descriptive or general ones? In other words, do singular representations share distinctive formal properties in addition to semantic ones – and if so, which? That is the question a psychological account of singular thought must answer.

Representations' formal properties cannot be determined from the armchair nor directly

observed. But they are inferable from characteristic psychological effects. To illustrate the relevant notion of 'effect', we borrow from Cummins (2010): Consider two mechanisms for multiplying. The first multiplies each digit of one factor by each digit of the other then adds the results. The second uses successive addition: it computes 3 x 3 as 3 + 3 + 3. The second mechanism exhibits the 'linearity effect': computational cost is a linear function of multiplier size. For example, 3 x 6 requires twice as many operations as 3 x 3, hence takes twice as long. The linearity effect is characteristic of the particular means through which the semantically-specified task of multiplication is accomplished. The effect is incidental to what the mechanism does, i.e., to the computation itself. Yet it hints towards how the mechanism works, i.e., the nature of the vehicles involved.

As the fact that the linearity effect can be diversely physically implemented suggests, the formal properties that interest psychology correspond not to representations' low-level physical (neurophysiological) characteristics, but rather to aspects of their causal-functional roles too fine-grained for task analysis to reveal. The fine-grained causal-functional properties that empirical research shows to be reliably shared by members of a psychological category constitute its signature properties (Carey, 2009). Such properties throw light on the deeper nature of the mechanisms which realize a psychological capacity, and so explain it (Cummins, 2010).

A plausible candidate psychological natural kind is a psychological capacity that exhibits many characteristic effects, leading to the discovery of a cluster of signature properties, and hinting towards a common set of underlying mechanisms.[5] Natural kinds are consequently "inductively deep" (Carey, 2009, p. 64): Members have projectible similarities which support generalizations. The signatures of a kind may therefore be used to determine its extension a posteriori. If certain effects or signatures are absent when a superficially similar task is performed, this suggests different psychological mechanisms are involved, and supports a 'split' in psychological classification.

Whether singular representations share a significant cluster of robustly projectible signatures, and so constitute a psychological natural kind, is an open question. An analogous question can be asked of mental files – representations whose theoretical-functional roles are also specified in broad functional-semantic terms by philosophers. As we interpret MFT, it claims that these two candidate psychological natural kinds, though distinct at the level of theoretical-functional role, turn out to exhibit many shared signature properties and effects, so that we are a posteriori justified in identifying them at the level of psychological realizers. The proposed reduction is local: files are (purportedly) what creatures like us think singularly with, even if others think their singular thoughts by other means. This hypothesis requires empirical support.

2.3 MFT as an empirical hypothesis concerning a candidate psychological natural kind

Recently, proponents of MFT[6] have appealed to psychological research on 'object-files'.[7] For

[5] Rather than having traditional essences, psychological natural kinds are 'homeostatic property clusters', groupings of entities which are suitable targets for scientific inquiry because they tend to share many interesting properties, which stably co-occur because of some empirically discoverable mechanism rather than by accident (Boyd, 1999; Ereshefsky & Reydon, 2015).

[6] While many philosophers appeal to some degree to psychology (Dickie, 2010; Jeshion, 2010, 2014; Montague, 2011; Recanati, 2010; Recanati, 2012, 2013; Sainsbury, 2005), others ignore psychological research entirely in their discussions of mental files (Crane, 2011; García-Carpintero, 2000; Hawthorne & Manley, 2012; Lawlor, 2001; Schroeter, 2007).

example, Jeshion (2010, p. 130) argues that the "essential singularity of mental files ... has its basis ... in the singularity of object files". Similarly, Recanati has recently defended MFT by claiming that he is "making an empirical hypothesis: that the object tracking system which exists in perception is used throughout cognition - even in high-level cognition" (Recanati, 2013, p. 212).

Object-files are thus supposed to furnish empirical justification not only for the reality, but also for the "essential singularity" (Jeshion) of *mental* files, which task analysis leaves open to doubt. However, as both Jeshion and Recanati remark, philosophers' notion of a mental file *extends* psychologists' notion of object-file. Even supposing *object*-files are somehow essentially singular, why should this property be shared by the broader class of *mental* files? As we reconstruct it, MFT relies on the following *projection argument:*

Mental files are a psychological natural kind. Object-files are a representative subspecies of the more inclusive kind mental files. Object-files have signature properties which make them distinctively singular and file-like, i.e., suited to the role of NMOPs of individuals. Signature properties of a representative subspecies of a psychological natural kind project across that kind. So, mental files, like object-files, have signature properties which make them distinctively singular and file-like. So, we are (empirically, defeasibly) justified in kind-identifying singular representations – the distinctive vehicles of singular thought – with mental files

To evaluate this argument, the exact relationship between object-files and mental files, and of singular thought to both of these, call for careful conceptual and empirical investigation.

3. Empirical research's contribution to MFT: Object-files, the standard model of files in psychology

In psychology, object-files are 'mid-level' visual representations: They operate between the low-level at which basic features such as edges, surfaces, textures, or contours are represented, regions are segregated, and figure-ground organization occurs, and the high-level at which categorization and recognition take place, and entities fall under semantically meaningful classifications such as 'wombat' (a general category) or 'Patrick' (a familiar individual). Between these extremes, object-files are created automatically when certain combinations of low-level features corresponding to visual 'objects' are encountered by visual 'input analyzers'. Files then 'stick' to these 'objects' as they move and change based primarily on their spatiotemporal characteristics, enabling us to track a limited number simply as *this* or *that* object, while storing and updating information about their features.

Let us now review some of the effects from which psychologists infer the existence and signature properties of object-files (Carey, 2009; Chen, 2012; Dickie, 2010; Scholl, 2001).

3.1 Object-based visual attention and memory effects in adults

Object-files are invoked by vision scientists to account for many effects suggesting that, in some sense, representations of 'objects' *per se* are 'units' of perception, at various stages:

[7] For example, Kahneman & Treisman (1984), Kahneman, Treisman, & Gibbs (1992), Scholl (2001), and see Section 4.

from before selective attention is allocated, to while attention is being deployed (governing its spread), to when information enters visual working memory (VWM[8]).

Object-files and attention

A notched circle interpretable as a full circle partially behind a square is hard to locate among an array of complete circles and squares. To find it, one must attend serially to the various figures in the display. However, a similar figure which is interpreted as in front of a square 'pops out' in visual search (Driver, Davis, Russell, Turatto, & Freeman, 2001; Rensink & Enns, 1998). A possible explanation is that, whenever possible, the visual system automatically fills in the notched circle so that it appears subjectively as a full cohesive shape, even before the subject attends to it, and even if this is detrimental to the search task. This suggests that *preattentive* – hence preconceptual (Dickie, 2010) – processes automatically carve up visual scenes into object-like units on the basis of low-level spatiotemporal cues, like cohesion. These objects then serve as possible targets for further processing, and for the assignment of object-files.

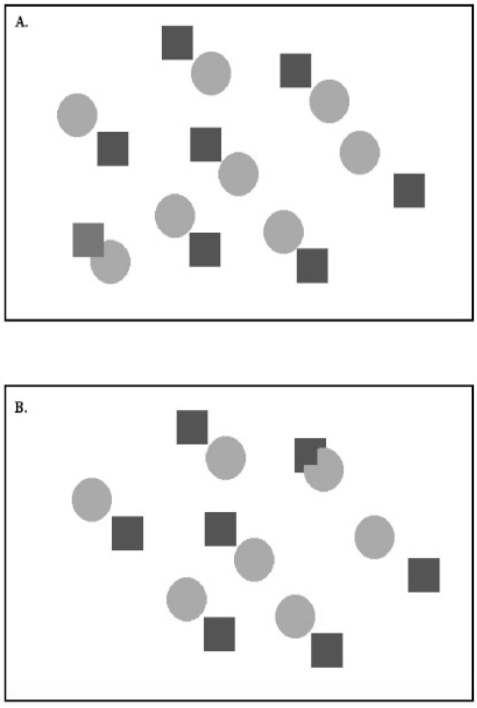

Illustration 1: Driver et al. 2001

Only a small number of objects can be pre-attentively indexed, and sustain later attentional processing, as suggested by the object-based effects exhibited by our capacity to visually track moving objects, which is studied using the multiple object tracking (MOT) paradigm (Pylyshyn, 2007; Scholl, 2009). In MOT, subjects see a screen containing 8-12 qualitatively identical items (e.g., same-sized white circles). A flashing subset gets selected as targets. Then the circles move around unpredictably for around ten seconds. When they stop, subjects have to point out the targets. Since targets and distractors are qualitatively similar and move, the targets must be tracked on the basis of their spatiotemporal trajectories, rather than features or

[8] VWM is memory in which visual information is actively maintained for ongoing tasks (Luck & Vogel, 2013).

fixed locations. Subjects can track only about four objects before performance sharply collapses. According to the standard model of object-files, this corresponds to the signature maximum number of files initially 'grabbed' preattentively by targets, and which can subsequently be maintained in parallel (though see below for controversies surrounding this interpretation).

Object-files influence how attention spreads within objects: attention moves faster to a target within the same static cohesive object than to an equidistant target outside of it (Egly, Driver, & Rafal, 1994; Moore, Stephens, & Hein, 2010) even when object boundaries are task-irrelevant (Chen & Cave, 2008).

Attention to objects also affects perception of *space*, both within objects and between objects. Dots appear farther apart inside the same object than when those dots do not appear on the same object (Vickery & Chun, 2010). Moreover, attended objects appear closer together than unattended objects (Liverence & Scholl, 2011). Allocation of attention *via* object-files thus seems to 'warp' how we perceive the surrounding space in which their referents are located.

Illustration 2: Vickery & Chun 2010

Object-files and visual working memory (VWM)

Object-based effects are also observed as attention allocated to objects influences further cognitive processing. Among these 'downstream' attentional effects are "object specific preview benefits" (OSPBs) (Kahneman et al., 1992; Noles, Scholl, & Mitroff, 2005): [9] In the object reviewing paradigm, subjects see a preview display in which visual information is associated with objects such as square boxes. One box in the middle of the display on top contains, e.g., the letter B. Another box in the middle of the display on the bottom contains an S. The letters disappear, the top box moves left, the bottom box moves right, then a letter reappears in one of the boxes. When asked to name the letter, subjects are faster when the same letter reappears inside the same box: For example, subjects are faster if B reappears in the left box rather than the right. Importantly, the locations in which the letters reappear in the final display are equidistant from those in which the letters are initially presented. This

[9] Analogous effects occur in the absence of objects, merely on the basis of textures (Ben-Shahar, Scholl, & Zucker, 2003). The precise signature of object-files may therefore not be mere preview benefits, but a specific *level* of preview benefits.

suggests that *objects as such* cause the priming effect, rather than locations.

According to the standard object-file model, OSPBs occur when (time-slices of) objects are assigned the same object-file on the basis of spatiotemporal factors, enabling faster access to entries already in the file. OSPBs thus are effects of object-based attention on retrieval of properties from visual memory.

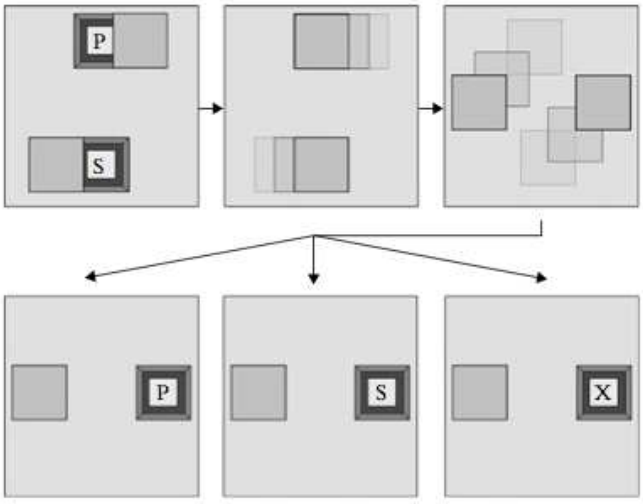

Illustration 3: OSPB, Carey, 2009

Object-files partially determine the capacity of visual working memory (VWM). Evidence comes mainly from change-detection experiments (Anderson & Awh, 2012; Luck & Vogel, 1997, 2013; Vogel, Woodman, & Luck, 2001). In these experiments, subjects see a sample array of items (e.g., colored shapes), and then after a short retention interval, a test array in which they must detect changes (e.g., a blue square has turned red). Subjects are capable of remembering only about four features at one time when these are distributed across an equal number of objects (four colors of four squares) before performance drops off sharply (five colors of five squares). Remarkably however, memory for four features distributed across four objects is nearly as good as memory for *sixteen* features distributed across four objects (four objects differing in terms of their colors, orientations, sizes and presence or absence of a "gap"). This suggests VWM stores information not in the form of a mere *list* of features, but contains a number of 'slots' for object representations in which features are *bound* together – files.[10] When the number of objects to be memorized exceeds the available number of slots, performance collapses. VWM capacity is thus limited by the number of files stored, somewhat independently of their total number of entries. (Though see section 4 for a competing explanation.)

The striking similarity between the nature and size of the capacity limits observed in MOT and in VWM has led most psychologists to consider that "object files and VWM [are] simply

[10] This VWM effect could also help distinguish the file model from a standard language of thought (LOT) model: To explain the effect, the LOT theorist has to specify – somewhat *ad hoc* -- that multiply tokening the same LOT term (a is F, a is G, a is H) is less costly in memory than tokening different terms in the same positions (a is F, b is G, c is H). Files seem to provide a more fitting metaphor for describing memory capacity.

two terms describing the same system" (Andrew Hollingworth & Rasmussen, 2010, p. 545)

3.2 More signature properties of object-files: Triggering and maintenance conditions

Object reviewing and MOT experiments suggest that the object-file system relies primarily on *spatiotemporal information* to determine when to open a new file and when to assign a pre-existing file to an element of the visual scene (Flombaum, Scholl, & Santos, 2009; Pylyshyn, 2007). Though qualitative features sometimes affect how vision 'decides' what counts as one and the same 'object' (Moore et al., 2010), the visual system reliably prioritizes lower-level spatiotemporal factors over high-level properties in computations of individuation and persistence. For example, in the tunnel effect, if the time it takes stimuli to successively appear on either side of an occluder is roughly the time it would take a single object to travel behind the occluder (so that the object appears 'temporally continuous'), and if the object appears to have followed a spatially continuous trajectory, viewers experience a single persisting object, even when the stimuli have different qualitative properties. For example, a yellow square is experienced as 'morphing' into a red circle inside the tunnel (Burke, 1952).

Analogously, in certain cases of apparent motion, subjects experience a succession of brief appearances of stationary objects as a single moving object on the basis of the timing and location of the flashes, regardless of their superficial features or apparent kinds. For example, flipping through a flipbook, subjects see a duck turn into a rabbit (Carey, 2009). Mitroff & Alvarez (2007) showed that OSPBs do not occur when objects in the preview and test display only share *features*: They must move on a spatiotemporally connected path (though see, Hollingworth & Franconeri, 2009). Plausibly, the various cases in which spatiotemporal information trumps competing featural or categorical data to cause the experience of a single persisting object all involve a common mechanism: the triggering and maintenance of object-files on the basis of such information. Supporting this hypothesis, Flombaum & Scholl (2006) showed that participants showed better performance on a color detection task for stimuli which respected spatiotemporal continuity compared to those which did not. Similarly, Odic, Roth, & Flombaum (2012) found that OSPBs line up with apparent motion effects.

Spatiotemporal priority, i.e., the fact that the visual system relies primarily on the spatiotemporal histories of objects to individuate them, is thus a "fundamental principle of object persistence" in vision (Flombaum et al., 2009), conformity to which is widely considered a signature property of object-files. Among the spatiotemporal constraints hardwired into the visual system's file-maintenance parameters, apart from spatiotemporal *continuity* (i.e. that objects move on connected, unobstructed paths) there is also the *principle of cohesion* (i.e. that objects move as connected and bounded wholes). When these principles are violated both MOT and OSPB are disrupted. For example, in MOT, subjects fail to track "stuffs" (such as water or sand) that move in characteristically substance-like ways (by pouring from one location to another) (Van Marle & Scholl, 2003). Likewise, when two objects merge into one, OSPB for only is observed, suggesting OSPB is sensitive to violations of cohesion (Mitroff, Scholl, & Wynn, 2004).

3.3 Object-based effects in infants

Infants have likely innate expectations about the behavior of physical objects, whose content is studied using violation of expectation and looking-time paradigms (Baillargeon, 1995; Carey, 2009; Spelke, 1990). This initial 'knowledge' of objects and their behaviors present many parallels with adults' object-files (Scholl & Leslie, 1999).

Signature capacity-limits

Infants' object representations are capacity-limited in similar ways to object-files: Infants are able to represent and maintain in memory only a small set of objects. They remember up to three objects stored in a container, but their performance collapses at more than four, which they do not even distinguish from more than two (Feigenson & Carey, 2005).

Triggering and maintenance conditions

In the individuation and tracking of objects, infants under 10-months privilege *spatiotemporal information* over property and kind information (Carey & Xu, 2001; Xu & Carey, 1996). For example, infants are shown a toy duck on top of a car. When a hand pulls on the duck's head, the car comes along with it. Unlike adults, infants are not surprised by this, suggesting that they fail to draw on kind-relevant qualitative differences to individuate static objects. In another of Carey and Xu's experiments, 10-month-olds see two stimuli appear successively from behind each side of an occluder, which differ in kind (a ball and a cup) or features (a red and a blue ball). This could be interpreted either as the ball changing (color or into a cup), or as two different objects successively appearing. Ten month olds expect one object behind the occluder, thus prioritizing spatiotemporal information in computing object persistence.

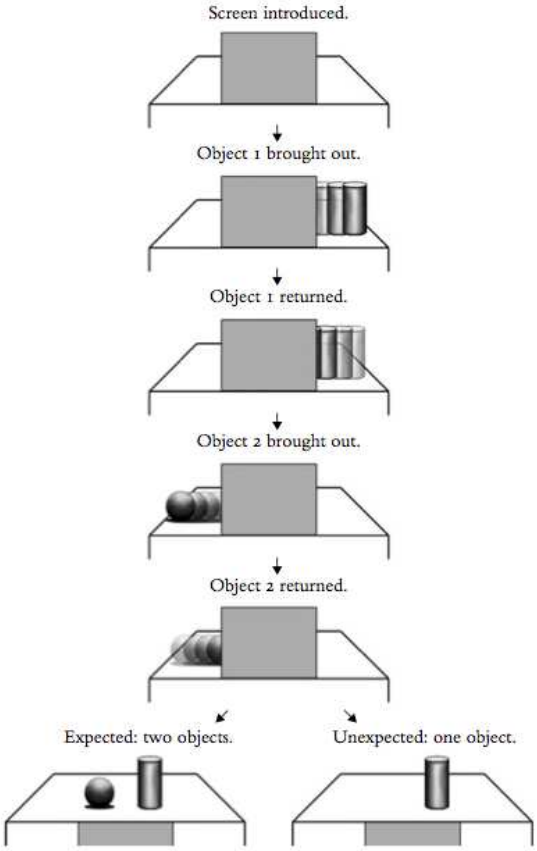

Illustration 4: Xu & Carey 1996, Carey 2009

Similarly to object-files, infants' object representations are maintained on the basis of principles such as *continuity* and *cohesion*. Infants are surprised when an object does not

appear in the middle gap between two occluders while traveling behind them (Luo & Baillargeon, 2005), thus violating spatiotemporal continuity. Violations of cohesion such as simply splitting a cracker in two disrupt infants' ability to track objects (Cheries, Mitroff, Wynn, & Scholl, 2008). Likewise, analogously to adult MOT, seeing piles of sand being poured from one location to another instead of moving as rigid unified wholes disrupts infants' ability to track small sets of objects (Huntley-Fenner, Carey, & Solimando, 2002). The object/substance distinction appears in infancy (Rips & Hespos, 2015) and is uninfluenced by cultural factors (Cacchione et al., 2014).

Summing up, according to the standard model of object-files in psychology, they constitute a psychological natural kind, a universal, domain-specific, innate, hardwired, and evolutionarily ancient *core system* (Carey, 2009; Spelke & Kinzler, 2007) for which there is converging evidence from independent sources. The main signatures of object-files are hypothesized to be as follows (Carey, 2009; Scholl, 2001):

1. Signature object-based effects on attention and memory, in particular object-based preview benefits and capacity limits.

2. Signature triggering and maintenance conditions, in particular, prioritization of *spatiotemporal* factors over featural cues in computations of individuation (initial opening of a file) and persistence (whether the same file 'sticks' diachronically to an encountered element).

3. Signature referential domain of 'Spelke objects', obeying principles of continuity, cohesion and contact.

4. Do object-files support MFT?

Object-files are often claimed to support a central *negative* claim of MFT. Supposedly, object-files empirically contradict a view of perception descriptivists assume, according to which perception represents only arrays of qualitative features scattered in space. On this view, individuals must be posited quasi-theoretically, as the hypothetical possessors of such-and-such immediately visible and epistemically transparent qualities, so object representation requires the sophisticated conceptual apparatus that goes with language and quantification. According to the object-file model, on the contrary, perception divides the world into objects prior to conceptualization, and offers them up as potential targets for thought. Object-files thus undercut the empirical basis for descriptivism in perception (Burge, 2010; Burnston & Cohen, 2012; Dickie, 2010; Pylyshyn, 2007).

In addition, object-files may be taken to support MFT's *positive* claims. Indeed, object-files have two main characteristics which arguably make them fit the bill for distinctively *singular* representations: First, unlike putative general files, object-files refer to *objects*. Even if the precise sense in which we should understand 'objects' in this context remains controversial, the function of object-files is (arguably) not to represent attributes or kinds (like wombat-hood or cuteness, nor even Spelke objecthood[11]). Second, unlike putative descriptive files, object-

[11] In the past, object-files have sometimes been presented as mental analogs of the general sortal concept 'object' (Xu, 2007). However, object-files on their own do not appear to support purely general or quantificational thoughts about objects as a category, such as the thought that *there are many objects in Ballarat.* We therefore grant MFT the now widespread view that object-files serve to represent token-objects, not the

files refer *nondescriptively.*[12] The signature triggering and maintenance conditions of object-files, based on principles like spatiotemporal priority, give *some* substance to the idea that files are nondescriptive. It thus seems plausible to appeal to object-files as a species of singular representations, and to claim they fill the role of NMOPs for *demonstrative* thoughts about objects.

However, before we fully endorse this important premise in MFT's projection argument, it is worth noting that the standard model of object-files is currently more controversial than philosophers acknowledge. This is doubly important to MFT. First, if *all* the characteristic effects associated with object-files are better explained without them, then the case for the psychological *reality* of mental files collapses. Secondly, if the standard picture of object-files is mistaken, object-files might exist, but fail to support MFT, by not being suitably singular or NMOP-like.

Are capacity limits file-based? A first area of ongoing debate concerns the nature and explanation of capacity limits in attention and memory. Understanding these limits simply in terms of a "magical number four" (Cowan, 2001) corresponding to the fixed maximal number of files available at one time is an oversimplification. Tracking capacity is affected by factors other than number of targets, such as objects' speed (Alvarez & Franconeri, 2007; Horowitz & Cohen, 2010; Howard & Holcombe, 2008), the distance between them (Franconeri, Jonathan, & Scimeca, 2010; Holcombe, Chen, & Howe, 2014; Scimeca & Franconeri, 2015), or the visual hemifield in which they appear (Alvarez & Cavanagh, 2005). This has led some psychologists to investigate alternative models of tracking capacity which do not posit an object-based bottleneck. For instance, some claim tracking is a serial process involving a combination of a high-capacity iconic memory and a roaming attentional spotlight (Holcombe & Chen, 2013; Oksama & Hyönä, 2004; Tripathy, Ogmen, & Narasimhan, 2011). Others propose that MOT capacity is determined by the flexible allocation of a continuous resource, which can variably spread across the entire scene, rather than a discrete number of files (Alvarez & Franconeri, 2007; Franconeri, Alvarez, & Cavanagh, 2013) or that it is entirely explained by spatial factors like crowding (Franconeri et al., 2010).

Similarly, VWM capacity does not depend solely on how many object-representations are deployed: We can remember fewer objects as their 'costliness' (featural complexity) increases. Again, it has been proposed that what accounts for VWM limits is *how much* of a continuous resource is being expended, not *how many* discrete file-slots are filled (Alvarez & Cavanagh, 2004; Steven L. Franconeri et al., 2013; Scimeca & Franconeri, 2015).

---

object-*type* as such (Carey, 2009; Marcus, 2001; Scholl & Leslie, 1999)

[12] Descriptivists could object that object-files depend on *some* property representations, namely the representation of objects' *spatiotemporal* properties, such as cohesion, rigidity, etc. As Burge (2010, p. 455) puts it, discussing Pylyshyn's FINST model: "Properties like spatial boundedness, spatial integrity, and continuity in motion are properties whose representation guides indexes for bodies." A radical anti-descriptivist response would be to maintain that the spatiotemporal properties which lead object-files to be opened and enable them to "stick" to visual elements diachronically are not – or need not – be *represented*: The instantiation of the appropriate properties simply *causally triggers* non-representational brain mechanisms which output and maintain object-files. The level of object-files would be where representational explanation bottoms out in *brute causal* explanation (Pylyshyn, 2007) A more moderate view would be to grant that there is some sense in which the lower-level properties the object-filing system responds to *are* represented, yet emphasize that they are only represented *within* the visual module: what makes an object the referent of an object-file is not its falling under concepts corresponding to Spelke principles, like the concept *cohesive*. For recent discussion, see for example (Orlandi, 2014).

Should this lead us to doubt the reality of object-files? Not necessarily. What matters for the object-file model, and consequently for MFT, is that attention and memory capacities be at least partly *defined in terms of objects*, not that the *number* of objects be absolutely fixed or that no other factors matter. For example, mixed models according to which these capacities depend *both* on allocation of some spreadable resource *and* the number of objects seem well supported (Alvarez & Cavanagh, 2004; Anderson & Awh, 2012; Hardman & Cowan, 2014; Luck & Vogel, 2013; Xu & Chun, 2009). VWM capacity could be the product of an interaction between object-slots and costliness of features, while tracking capacity could depend on both the number of targets and their speed, distance, or other factors. For example, in the case of VWM, it may be that while in many circumstances the number of files available is four, this number changes depending on how difficult the features inside the files are to remember. For complex features, the number of available files may drop to two. Nevertheless, if the object limit is two instead of three, participants will now be better at remembering three features per object for two objects than two features per object for three objects. At any set level of featural complexity, there are still *some* object-based limits on memory (this is an empirical testable hypothesis). Object-based limits in tracking may work in an analogous fashion (this is an area for future experimental inquiry). This is enough to support *some* version of the object-file model, if not the oversimplified one in which there are *exactly* four files.

Are individuals per se being tracked? Some psychologists deny that MOT involves the representation of objects *as individuals.* This objection is especially relevant to the claim that object-files fill the demonstrative NMOP-role. The source of the objection is the finding that if individual *names* (or other identifying characteristics) are assigned to each of the targets before they start to move, subjects find it near impossible, when they stop, to say *which* target has which name. For example, supposing one of the targets was initially baptized 'Pat', subjects cannot single out Pat at the end of the task. Subjects only know whether a given object is *a* target, not w*hich* target *it* is (Pylyshyn, 2004; Scholl, 2009).

Scholl (2009, p. 57) argues this should lead us to question "one of the key assumptions about MOT … that each target object is being tracked as a distinct individual: during tracking one is keeping track of *this* target, *that* target, and *that* target as each moves about the display". Thus, whereas Pylyshyn takes MOT to show that object-files are linked to objects thanks to some pre-attentive mechanism analogous to pointing fingers, Scholl (ibid.) retorts that this "is essentially equivalent to tracking two objects by continually pointing to one with each index finger, but then later having no idea which object you were initially pointing to with your left index finger!". In philosophical terms, this would correspond to either a violation of Frege's Constraint, or to a violation of the Transparency Constraint on MOPs: either representing an object under the same/different files would not amount to representing it as the same/different object from the subject's perspective, or subjects would be incapable of knowing, purely introspectively, whether they are representing an object under the same or different MOPs/files (at least diachronically[13]). On Scholl's alternative picture MOT simply requires ordinary *multifocal attention*, which need not distinguish its different foci as separate *individuals*.

[13] It is sometimes claimed that transparency only holds synchronically. However, one of the central historical philosophical motivations for files was clearly to account for *continued* belief (Perry, 1980). Furthermore, even philosophers who reject diachronic transparency maintain that file-identity is transparent during a single exercise *of the same tracking capacity* (Campbell, 1987; Recanati, 2012) which lack of transparency during a single episode of MOT contradicts.

Scholl's objection to Pylyshyn's model seriously threatens MFT. Yet several points are worth noting in response. First, even if Scholl's argument tells against Pylyshyn's claim that tracking involves pre-attentive indexing, it actually does not matter so much to the object-files model if tracking requires *attention* after all, so long as it does not require the deployment of descriptive conceptual resources stored in long-term memory – which does not seem to be the case. Second, it is unclear why the incapacity to remember, e.g., a *name* associated with a target at the end of a MOT experiment means that the object is not being tracked as an *individual*: in this context, names function as yet another predicative *property* (entry) inside files, which we know subjects are poor at remembering during tracking anyway (Bahrami, 2003). Roughly, an object-file based demonstrative thought would have the structure '*that* is Pat', with 'is Pat' serving as a file-entry. Our incapacity to re-access individuating properties may pose a threat to the claim that *entries* inside object-files are stably stored throughout MOT, but it does not refute the claim that tracking involves *singular* reference. Thirdly, variants on MOT have been increasingly studied to test competing models of tracking. Among these is Multiple Identity Tracking (MIT), in which the task is not simply to locate several indistinguishable targets but rather distinct individuals, e.g., the (only) wombat among animals of other species (Oksama & Hyönä, 2008; Pinto, Howe, Cohen, & Horowitz, 2010). Results suggest that subjects *can* keep track of the individualizing characteristics of objects in more ecologically valid situations. In sum, for the purpose of the standard object-file model, what really matters is that subjects have the capacity to track at least *some* individuals without relying on conceptualizations of their properties. And that much remains empirically plausible, regardless of exactly how we explain standard MOT.

Are object-files transparent? Although ordinary MOT might not show a violation of the Transparency Constraint, other empirical results suggest that, unlike MOPs, object-files are not always present transparently to consciousness.[14] In an experiment (Mitroff, Scholl, & Wynn, 2005), subjects saw an ambiguous display, in which objects are perceived either as streaming through or bouncing off each other. When the display was manipulated so subjects consciously saw streaming, OSPBs were still observed in the opposite bouncing direction. Mitroff et al. (2005, p. 67) conclude that "conscious percepts of 'which went where' in dynamic ambiguous displays can diverge from the mapping computed by the object-file system." Thus, object-file identity may not always be available to subjects consciously and transparently. Converging evidence is provided by (Norman, Heywood, & Kentridge, 2013): Adapting Egly et al. (1994)'s demonstration of object-based attention, these authors showed that targets are processed faster within the boundaries of an object of which a signal-detection task indicates subjects are unconscious.

Though the fact that object-files do not always line up with conscious percepts should not prevent us from tying object-files to conscious perception in *most* cases, it nevertheless makes it doubtful whether *transparency* is among object-files' signature properties as a kind. More broadly, file-based explanations and Fregean ones involving MOPs have different aims: While the former focus on capturing subjects' rational personal-level perspectives, the latter use (unconscious) effects to identify subpersonal mechanisms.[15] It is therefore not surprising that

[14] For example, Peacocke (2001, p. 253) writes: "the way in which some thing ... is given in the nonconceptual content of an experience is something which contributes to what it is like to have that experience. These ways which feature in nonconceptual content are then at the conscious, personal level, and are not merely subpersonal. As features of the subjective experience, their presence can entitle a thinker to make a particular judgement, or to form a certain belief." See also (Campbell, 2011)

[15] As Hollingworth & Rasmussen (2010, p. 543) put it: "it is important to consider that the original object reviewing paradigm of Kahneman et al. (1992) did not probe conscious perception; all that was measured was priming. Thus, claims regarding the updating of memory with motion need not be inextricably bound

object-files should violate Fregean constraints. Nevertheless, defenders of MFT can respond by pointing out that transparency is best construed as a *normative* claim, and that violations of transparency are *abnormal.* If object-files *normally* line up with perceptual MOPs, that might be good enough for MFT: on a view of natural kinds as homeostatic property clusters, even highly projectible properties may be absent in sufficiently deviant contexts. Whether or not transparency violations are indeed exceptional is an empirical issue.

In summary, the premise that object-files *exist* and are *singular* has evidence in its favour. While object-files might not behave exactly like NMOP-role fillers, the mismatches are (arguably) tolerable: *a posteriori* reductions are seldom perfectly smooth. Other parts of MFT's projection argument remain problematic. MFT faces *the problem of range*.

5. The problem of range for MFT

Mental files cannot be modeled too closely on object-files, or they fail to account for *all* singular thoughts (for reasons given below). But if mental files are made to depart too much from object-files, this guts the notion of empirical substance, and weakens the grounds for projecting properties of object-files to mental files more generally.

To avoid the first horn of this dilemma, the notion of 'mental file' must be extensive enough to account for *all* singular thoughts. To avoid the second, it must be sufficiently restricted to bear a clear kinship to 'object-files'. This is the *problem of range* for MFT. It manifests in two ways.

5.1 The problem of referential range: Object-files are for objects

The first aspect of the problem concerns *referential* (or *categorial*) range. One of the signatures of *object*-files is that they target Spelke *objects* (henceforth 'objects')[16]. Yet the entities we think singularly about ('individuals') are not all objects, but a motley crew – from deities to cake-delivering businesses (Jeshion, 2010; Kim, 1977). Many fail to conform to Spelke principles or the stereotype of "middle-sized dry goods" (Austin, 1964; Bloom, 2002; Hansen & Rey, forthcoming). It is for example unclear how object-files could account for thoughts about the following: animate agents (Patrick, a snake), undetached body parts (a snout), non-existent and fictional entities (Santa Claus), events (the writing of this chapter), locations (Ballarat), times (now), situations (the actual world), 'minor' entities (Patrick's shadow), ephemera (bubbles), very large or small individuals (the Great Barrier Reef), surfaces or backgrounds (the sky), collections or groups (a wisdom of wombats), abstract individuals (the number seven, the Wombat Preservation Society), paths and trajectories (the route from here to Ballarat), mental states (this pain), oneself, and objects made up of disconnected parts (disassembled watches).

The problem of *referential range* is thus the following: The signature domain-specificity of object-files supports MFT to the extent that it gives empirical substance to the claim that *real* files are for representing immediately perceptible *objects* as opposed to attributes or kinds. But this blessing is also a curse. Object-files' referential domain is too narrow. The challenge is to account for singular thoughts about individuals that are not immediately perceptible or not *objects*, hence not covered by object-files, without ceasing to employ a notion of 'file'

---

with claims regarding conscious perception."

[16] "[T]he capacities infants have are not tracking mechanisms but rather *object*-tracking mechanisms." (Bloom, 1998, p. 67)

which picks out representations of the same psychological natural kind.

5.2 Extending files' referential range: The horizontal extension of files (within perception)

To extend files *horizontally* is to claim that there are more species of files at the pre-conceptual, mid-level perception than merely object-files. More precisely, the claim is that there are core representations which closely resemble object-files *formally*: They cause similar downstream effects on memory or attention, and have signature triggering and maintenance conditions, corresponding to specific ways of singling out individuals independently of higher-level descriptive classification. Yet these representations are not *object*-files, and their dedicated domain is not that of *objects*. These additional species of file could help MFT explain demonstrative thoughts about immediately perceptible entities that are not objects.

Superficially, horizontal extension of files resembles MFT's usual multiplication of files. But it differs in substance and method. In substance, because it is compatible with files only existing in perception; in method, because it is grounded on empirical evidence that other representations share object-files' signatures. For reasons of space, we present just one illustrative application of this strategy: We argue that mid-level perception includes *event-files*, which ground demonstrative thinking about (some) particular visual events *as events*.

5.3 Event-files

To represent an event, one usually represents an object participating in it.[17] It does not follow that event representation requires descriptive conceptualization or language. Indeed, recent empirical evidence suggests that just as there are representations of individual objects in core cognition, there are also representations of individual events[18]: event-files (Hommel, 1998[19], 2004; Shipley, 2008) "The ongoing experience of events ending and beginning would then correspond to the opening and closing of such files" (Shipley, 2008, p. 21) Defenders of MFT have so far ignored this species of file.

Imagine a situation where someone grabs a pen, clicks it, writes, and puts it down. Although the stretch of time in which this happens is continuous, we tend to view this scenario in terms of four discrete units (subsequently combinable into one larger one). In an early experiment, Newtson and Engquist (1976) had participants view videos and press a button wherever they felt like "one meaningful unit of activity has ended and another one has begun." Participants generally agreed as to where event boundaries lie, suggesting some common mechanism could underlie event segmentation. Further research suggests this mechanism is *perceptual*, by showing that event segmentation is largely automatic, and has an important role in on-line processes of visual memory and attention (Radvansky & Zacks, 2011).

Event-files & attention. Event boundaries, like those of objects, play a major – though complicated – role in structuring attention. While the adult visual system seems better at detecting small interruptions at event boundaries (Newtson & Engquist, 1976), subjects are worse at detecting visual probes appearing at event boundaries (Huff, Papenmeier, & Zacks,

[17] Yet one can see a shimmering or a lightening, without there being some *object* that shimmers or lightens.
[18] For the purpose of this section, we treat actions as a subspecies of events.
[19] Hommel appears to be the originator of the concept. However, he uses it differently than we do: Very roughly, he seems to consider an event-file to be simply an object-file which stores action-relevant features (like affordances).

2012).

Event-files and memory. Further supporting the analogy between event representations and object-files, individual event representations structure automatic processes of VWM. For example, memory for scenes or moments from an event improves at the boundaries (Swallow, Zacks, & Abrams, 2009) and when observers view videos depicting human interactions with simple everyday objects, memory for items from an ongoing video is superior when no event boundary occurs between presentation and test compared to a case in which such a boundary does occur, even when the temporal delay between stimulus and test is controlled for (Swallow et al., 2009). These findings indicate that event representations are on-going constructs in working memory whose contents are 'cleared-out' once a segmentation point is encountered.

Event representations determine the capacity to remember distinct events in a manner strikingly reminiscent of the signature capacity limits of object-files. Wood (2007) presented observers with between one and five brief sequences depicting continuous human motions (e.g. a person raising an arm). After exposure, participants were shown a second sequence in which all the actions were identical to those presented during exposure or in which one of the actions had changed (e.g. a person raising a leg instead of raising an arm). The participants' task was to indicate whether the second set of actions was the same or different from the first.

Wood found that we can only store representations of about two or three separate events in working memory at once. Importantly, these memory limitations cannot be explained by overall duration of activity, overall amount of movement, or other working memory constraints like *object*-based limits (thus clearly distinguishing these representations from object-files). Strikingly, observers were as good at remembering three properties distributed across three events (e.g. each event having a unique action category) as they were at remembering nine properties distributed across nine events (e.g. each event having a duration, category, and side of the body). But they were worse at remembering four properties for four events than nine properties for three events. These results strongly suggest that the capacity for visual memory is (at least partially) measured by representations which it makes excellent sense to conceive of as 'event-files', given the clear analogy with the VWM effects studied by Luck & Vogel.

Triggering & maintenance conditions. Since event representations, like object-files, are activated reflexively during perception, it is natural to ask which specific visual cues trigger their activation, and whether these mirror those in object perception.

Many such parallels exist. While objects are perceived as cohesive units in *space* with clear boundaries defined by object contours, events are perceived as cohesive units in *time*with clear boundaries defined by moments of salient change (Zacks & Swallow, 2007).

Recent evidence shows that *cohesion* constraints on event representations induce false memories for events due to processes that mirror object-based amodal completion. Thus Strickland & Keil (2011) showed participants videos depicting a causal event (e.g., a person kicking a soccer ball) and a series of pictures after the event. In the crucial conditions, the video was edited in such a way that the moment of contact between foot and ball was missing. Participants nevertheless reported having seen this contact when it was heavily implied by subsequent video footage (e.g. when the participant viewed the resulting flight of the ball) but not when contact was not implied (e.g. when the participant viewed irrelevant footage from

the same scene). It also appears that preverbal infants make such causal bridging inferences. For example, pre-verbal infants around 10-months spontaneously infer the unseen causes of visible effects even for highly unfamiliar events (Saxe, Tenenbaum, & Carey, 2005).

Pursuing the analogy to objects, *continuity* also appears to be an important triggering condition for event representations. In Magliano and Zack's (2011) FMRI study, the authors contrasted videos which were edited in ways that involved (a) no discontinuities, (b) spatial/temporal discontinuities which maintained event continuity or (c) event discontinuities which maintained spatial/temporal continuity. Event discontinuities disrupted segmentation behavior more than spatial/temporal discontinuities, and event discontinuities provoked a unique pattern of neural responses which differed substantially from videos containing no discontinuity, and videos containing only spatial/temporal discontinuities.

In summary, empirical evidence suggests that the visual system automatically parses the flow of time into discrete representations of individual events, that then go on to structure attention and memory analogously to object-files. There is thus reason to posit event-files, another subspecies of file in core cognition alongside object-files, which could account for *demonstrative* thoughts about certain basic visual events.

Furthering the horizontal extension? We have proposed that there are at least two file-systems within core cognition, belonging to the same psychological natural kind but underlying distinct modes of singular reference to different domains: object-files and event-files. The problem of referential range is far from solved. But perhaps the same horizontal extension strategy could be pursued for other categories of individuals.[20] For example, (Murez & Smortchkova, 2014) argue for the existence of 'agent-files', which account for demonstrative thoughts about animates agents (which they argue are cognitively distinct from objects). (Tsompanidis, 2015) argues for files that represent moments of time.[21] Whether mid-level perception contains these or yet other species of files is an open issue. We turn to another aspect of the problem of range, not solved by the postulation of files for non-objects in perception.

6. The problem of cognitive range: Object-files are not concepts

Singular thoughts are not all *demonstrative*. They also include thoughts of at least the following three varieties:

Memory-based thoughts. Sitting at your desk, you think freely about Patrick. Yet nothing impinging your senses relates to him. One can thus form conscious, singular *occurrent judgements* simply by activating singular *standing beliefs* stored in long-term memory.

[20] Boyer (1998, pp. 67-68) suggests that "domain specificity could start at a low level of processing, that of "tracking" same-substance cues. Representing two faces as the same person with different emotions would activate cognitive resources entirely different from those involved in recognizing two sheep as same kind … Note that this does not imply a return to the view that having a concept is having a description. All that is required is that (1) objects are assigned to ontological domains on the basis of perceptual cues … (2) this makes the system attend to or expect "same substance" information in a way that is specific to the domain."

[21] More radically, Millikan (1998, 2000) argues that files are involved in nondescriptive thoughts about all "substances", a category that includes not only individuals of various sorts, but also stuffs (like milk) and "real kinds" (like wombats). However, empirical evidence for her view is somewhat lacking.

Communication-based thoughts. You entertain singular thoughts about Patrick despite having only ever read about him. This illustrates the capacity to form singular thoughts about individuals simply by understanding utterances or texts about them.

Recognitional thoughts. You repeatedly recognize some individual, encountered under a variety of perspectives, as Patrick – one and the same individual across various contexts.

Such thoughts are to various degrees perceptually *detached*. By contrast, object-files are *encapsulated*.[22] Consider again the tunnel effect. One might experience a ball as changing into a cup, while judging that no object actually transformed. Even when this squarely contradicts our firmly-held rational beliefs, object-files make us "irresistibly perceive" (Flombaum & Scholl, 2006, p. 840) objects as the same, so long as they obey certain constraints, which are similarly hardwired into the visual systems of newborn chickens (Rugani, Fontanari, Simoni, Regolin, & Vallortigara, 2009)[23].

Like domain-specificity, this signature of object-files cuts both ways for MFT. On the one hand, it supports the claim that *real* files are *nondescriptive*: The properties which trigger and maintain them are only represented within dedicated low-level 'input-analyzers', if at all. On the other hand, sensitivity to logical reasoning or theoretical beliefs is one of the markers of singular *concepts* and MOPs.[24] MFT claims files explain the cognitive significance of singular terms in natural language, such as indexicals or proper names (Lockwood 1974, Bach 1987, Recanati 2012, Evans 1982, Forbes 1989, Jeshion 2009, 2010). To fill these roles, files must feature in higher-level, language-based cognition, where rational or logical norms apply. This gives rise to the problem of *cognitive* (or conceptual) *range:* (Purported) *conceptual* files are unlike object-files in many ways. Notions of 'file' lumping them together risk being *too* inclusive, not only to cut psychological reality at the joints, but also to demarcate singular from general or descriptive thought.

Defenders of MFT see this less as a problem than as an opportunity to distinguish multiple *subspecies* of files: For example, Recanati (2012) posits re-deployable 'recognitional files' to account for recognitional thoughts, and 'encyclopedic files', to account for memory-based thoughts. Jeshion (2010) posits files based on 'mental names'. Such files' *continuity* with object-files is emphasized. Jeshion (2010, 2014) claims object-files smoothly develop ontogenetically into mental files based on 'mental names', and Recanati (2013: REF) hypothesizes (recall) that "the object tracking system which exists in perception is used throughout cognition - even in high-level cognition".

Proponents of MFT have not confronted a competing model psychologists take seriously, which stresses the *discontinuity* between the object-file system, and what is considered a distinct higher-level system for individuating and tracking individuals based on beliefs about properties and kinds (Carey, 2009; Carey & Xu, 2001). This second system is what one would rely on to judge, for example, that although a ball appears to have turned into a cup behind an occluder, these objects are really distinct. According to Carey & Xu (2001, p. 203) "once this second system of kind-based object individuation has become available, it creates the

[22] The object-file system does not take subjects' inferential reasoning or theoretical beliefs as inputs.

[23] The same point applies to synchronous object individuation. For example, no matter how convinced someone is by the metaphysical principle according to which any mereological sum of objects whatsoever composes another object, their object-file system will not assign a single file to, e.g., the sum of Patrick + Eiffel Tower.

[24] As Prosser (2005, p. 370) puts it: "What is essential to MOPs is that they are individuated in such a way as to make maximum rational sense of the thinking subject."

representations that articulate thought. That is, it preempts object file representations in our experiences of the world." On this model, files occupy a circumscribed area of cognition. Rather than outgrow core cognition during development, they are superseded by representations of a different kind.[25]

To decide between these competing models, we empirically investigate the relevant empirical literature for MFT's hypothesis according to which files constitute a unified psychological natural kind, extending beyond mid-level perception into higher-level singular thought, i.e., the *vertical extension* of files.

7. Vertical extension of files (beyond perception)

Two types of empirical arguments for the vertical extension of files are worth considering. Firstly, there could be *direct* evidence that representations in higher-level cognition have analogous signatures to object-files. Secondly, there could be more *indirect* evidence, such as developmental or evolutionary connections between them.

7.1 Direct evidence

Signature capacity limits

Capacity limits are a signature of object-files that philosophers mostly ignore in characterizing mental files.[26] Could an empirical case be made for files in higher-level cognition based on capacity limits?

Kamp (1984) and Recanati (2005) speculate that information used to interpret context-sensitive expressions such as indexicals is available in the same format regardless whether it is acquired through previous discourse or perception. Recanati hints at an argument for this hypothesis based on signature capacity limits. He gives the following case:

> Yesterday, my brother talked to the policeman about the burglar we saw. He told him he thought he had escaped, but the policeman would not believe him, arguing that someone was awake, and he would have seen the burglar if he had left.

This is supposed to suggest that *verbal* working memory – and more specifically, the capacity to keep track of the identities of highly salient singular referents of anaphoric pronouns (underlined) – displays the same signature limit as *visual* working memory/attention found in MOT.

---

[25] It should be noted that object-files are described as 'concepts' by some psychologists (Carey, 2009; Spelke, 1990), roughly because they are multimodal (Jordan, Clark, & Mitroff, 2010) and available as input to action planning. But the disagreement with the discontinuist view is only apparent, since this notion of 'concept' is less demanding than the way in which 'concept' is used in philosophy, roughly for representations that are inferentially promiscuous and MOP-like. The point that core representations are distinct from the constituents of higher-level beliefs, including folk theories, is indeed central to the work of the same psychologists who describe them as 'concepts'. For example, Carey & Spelke (1996, p. 519) note that: "theories are always open to revision, including radical revision including conceptual change or even abandonment. Core systems, in contrast, are elaborated but not revised."

[26] For example, Fodor (2008, p. 94) writes: "Think of your head as containing (inter alia) an arbitrarily large filing cabinet, which can in turn contain an arbitrarily large set of files." Although Fodor references Kahneman, Treisman and Pylyshyn as the source of this usage of 'file', it is not clear what indefinitely storable and numerous 'files' have to do with object-files, which are essentially limited in number and short-lived (Noles, Scholl, & Mitroff, 2005).

A similar view has been independently proposed in psychology by Cowan (2001) who claims that the 'magical number four' sometimes considered a signature of object-files actually reflects a *central* limit on attention. Another theorist who has defended a similar view is Hurford (2003) who draws a parallel between MOT capacity limits and linguistic deixis, noting that few languages have more than five contrasting deictics. According to such views, files are in central cognition, rather than confined to specialized core systems.

Taken literally, however, the hypothesis according to which files underlie a *unified* system for referent-tracking in perception and discourse predicts that holding verbal items in WM interferes with visual MOT or WM tasks. This prediction is not corroborated: Luck & Vogel (1997) found no effect of verbal load on their visual WM task. And tracking four objects in MOT while remembering four verbally presented items is trivial (Scholl & Xu, 2001).

In response to this objection, the hypothesis can be interpreted less narrowly: There could be several *separate* filing systems – verbal, visual, and modality unspecific, perhaps – with analogous capacity-limits. Scholl and Xu (2001) raise a conceptual challenge to this proposal: The central limit posited by Cowan and others, supposing it exists, is defined over associative *chunks*, a notion far less constrained than 'file'. More generally, even if capacity-limits in higher-level cognition were shown to be *numerically* similar to limits in core object cognition, evidence that these limits are defined in terms of *file*s and not other information structures is needed.

One area where this challenge might be taken up is semantic memory, which exhibits the 'fan effect' (Anderson & Reder, 1999): the more (recently learned) facts one knows involving a concept (associations 'fanning off' from it), the harder it is to retrieve any one of these facts. More significant for our purposes is the *differential* fan effect (Radvansky, 1999): the strength of the fan effect depends on the concepts and facts involved. For example, it is harder to retrieve facts about a single object in many different large locations ('the palm is in the hotel', 'the palm is in the library', 'the palm is in the school'), than facts about a single location with different objects ('the palm is in the hotel', 'the phone is in the hotel', 'the plate is in the hotel'). For pairs of *small* locations (that can contain just one person) and people, the effect is reversed: it is easier to remember one person in several locations, than several people in one location. Radvansky interprets these results in terms of 'situation models': integrated representations of token individuals and pieces of information about them. If the pieces of information are consistent with a single situation, they are integrated into a unified model. The fan effect occurs when information is distributed across multiple models, each of which must be separately accessed for retrieval.

Situation models differ from, e.g., Recanati's files, to the extent that what determines integration of pieces of information is their content, not the ER channel they issue from. Still, the resemblance of 'situation models' to mental files in semantic memory, as opposed to an unconstrained notion of 'chunk', should be clear: Like files, what ties 'situation models' together is plausibly *copredication*. The differential fan effect thus supports a picture according to which it is harder to access information distributed across multiple files rather than co-filed. Furthermore, since the fan effect occurs for a variety of non-object individuals (such as people, situations, and locations), it hints towards the psychological reality of file-like structures for such individuals.

Nevertheless, since the differential fan effect has no *strict* analog in mid-level perception, it is unclear how much such files resemble *object*-files. In ongoing experiments, we investigate

whether semantic memory exhibits effects more closely analogous to those Luck & Vogel and colleagues found in VWM for objects, and Woods found for events. In these experiments, subjects perform a change detection task akin to the Luck & Vogel paradigm, except that the information to be retained is semantic rather than visual: Participants have to memorize up to three pieces of information about a varying number of people, and then detect changes – perceptually detectable changes in the displays are irrelevant. We have so far not found any characteristic 'filing effect'. Performance is determined solely by total number of features to be remembered: when the number of features increases, performance decreases linearly.

Signature triggering and maintenance conditions

Another signature of object-files are their signature triggering and maintenance conditions, based on such factors as spatiotemporal priority, and criteria of cohesion, continuity and contact.

At first glance, these signatures might appear not to apply to higher-level concepts of individuals. Consider a situation in which you identify some creature *as Patrick* as the result of effortful abductive inference, e.g., "I am in Ballarat. That is an old and large wombat. Everyone is calling it 'wombassador'. So it must be *Patrick*". In terms of the file metaphor: what triggers a certain putative file, and makes you apply it to an individual, is that it seems to you upon reflection to match the file's entries, not that preconceptual input-analyzers spit the file out automatically on the basis of spatiotemporal factors. Clearly, *access* to such a (hypothetical) file is descriptive: it depends on a "theory-based sustaining mechanism" (Margolis, 1998, p. 354), not an ER-based one. As previously noted, Carey and Xu differentiate concepts of individuals which are thus descriptively triggered and applied from files.

In response, defenders of MFT would no doubt remark, firstly, that the fact that a file is descriptively *addressed* does not necessarily make the file non-singular in other senses: its reference, individuation, and maybe even entries, could still be determined nondescriptively. Secondly, there could still be *some* sense in which the file is applied on the basis of predominantly spatiotemporal factors. Scholl (2007) argues that the sorts of intuitions relevant to metaphysical debates about object persistence reflect, more or less directly, the signature triggering and maintenance conditions of object-files: what we judge to be one and the same object at the fully conceptual level are entities we believe to be of the sort that object-files could track, i.e., that move on spatiotemporally continous paths, etc.

The problem is that, even if there is an obvious resemblance between our folk criteria of object-persistence and the maintenance conditions of object-files, it is equally clear that our mature criteria of individual-persistence depart significantly from these. Rips and colleagues (Blok, Newman, & Rips, 2005; Rips, 2011) have not found spatiotemporal factors to be decisive in persistence judgments at the fully conceptual level. Well-known intuitions about non-spatio-temporally continuous *persons, artifacts,* or *abstracta* in the philosophical literature also suggest this (Gallois, 2008). For example, many of our intuitions about personal identity seem to have more to do with hidden psychological essences, than spatiotemporal factors. An interesting case is noted by Gelman & Bloom (2008): to be elected the 14th Dalai Lama, a boy must choose between qualitatively identical possessions of his previous incarnation, which are thought to be imbued with the essence of the 13th Dalai Lama. Such singular essence-based tracking appears to operate according to different psychological principles than the spatiotemporally-based tracking of the object-file system.

Signature domain-specificity

As the previous subsection suggests, one of the apparently striking differences between object-files and higher-level singular concepts has to do with *domain-specificity*: moving beyond concrete Spelke objects, adults acquire a *formal* conception of individuals, corresponding to (roughly) whatever singular terms pick out (Bloom, 2002; Casati, 2004).

It remains tempting to see *some* connection between the capacity to master singular terms, and the capacity to perceptually track objects. Gentner (1982) influentially proposed that children are biased to interpret novel words as names of whole-objects. If someone points towards Patrick, and says 'Patrick', a word the child has never heard, she tends to interpret it as referring to *Patrick*, and not his tail, one of his properties, some event, etc. Indeed, *names* of objects are frequent among children's first words (Hall, 2009). Although we eventually give names to all sorts of entities, it might seem that, at some basic level, the proper-nameable entities match the dedicated domain of object-files. There could thus remain a trace of object-files' signature domain-specificity at the heart of our linguistic capacity.

Among proponents of MFT, Jeshion (2009, 2010) puts weight on the connection between the domains of files and proper names. However, rather than conceive of proper nameable entities as objects, based on object-file research, she infers a (novel) signature property of files from what she takes to be a signature property of names: Jeshion (2010, p. 136) hypothesizes that files obey a 'significance condition', according to which "a mental file is initiated on an individual only if that individual is significant to the agent with respect to her plans, projects, affective states, motivations." Jeshion sees this purported signature property of files as evidenced by our naming practices: We only give proper names to *significant* individuals.

Thus we have a potential argument for vertical extension of files in which domain-specificity figures: Files preferentially target the domain of significant individuals, like names. However, this argument is problematic.

The fact that people are proper-nameable *par excellence* (Hall, 2009) might seem to support Jeshion's view: people are significant. But projecting mental files' significance to object-files is empirically unjustified. Despite Jeshion's appeal to object-files as close ontogenetic precursors of mental files, there is no empirical evidence that they obey the 'significance condition'. On the contrary: Krøjgaard (2000) investigates whether significance matters for object-files using a modified version of the Carey & Xu paradigm: One object is the child's favorite toy, hence highly significant, whereas the other is unfamiliar. Results suggest "it makes no difference for the infants' reactions whether the objects that disappear behind a screen are novel or significant" (Krøjgaard, 2000, p. 181). This is unsurprising, if the object-file system is an encapsulated system which does not care about anything much apart from spatiotemporal factors. But this result is unexpected on Jeshion's picture, where files resemble names by obeying the significance condition.[27]

In sum, direct evidence that representations in higher-level cognition share signatures of object-files is currently mixed. In the absence of clear-cut evidence for vertical extension, in the following section we consider more indirect evidence.

[27] Jeshion might respond that while the internal computations of the object-file system don't obey the significance condition, the mere fact that an object-file is tokened for an individual *makes* it significant to the subject. However, this seems hard to square with the possibility of unconscious object-files.

7.2. Indirect evidence for vertical extension?

Developmental and evolutionary considerations have been adduced in favor of files' extension beyond core cognition. For example, Recanati (2012, 2013) justifies vertical extension by appealing to files' acquisition of "derived functions": Files whose primary function is to refer to perceived individuals acquire the far broader functions associated with discourse reference. Similarly, Jeshion (2010, p. 135) proposes that FINSTs/object-files are "coupled" with linguistic demonstratives, and "mental demonstratives, construed as a type, come to function as mental stand-ins for FINSTs. They develop so as to function constitutively as abstract singular referring devices by means of which we think singularly about individuals."[28] After mentioning one source of data that could help make these proposals more than attractive 'just-so' stories, we conclude by noting some challenges.

The evolutionary story MFT tells to justify files' vertical extension connects attentive tracking to pointing, and then pointing to linguistic deixis and, ultimately, discourse reference. Leaving aside temporarily the first step of this process (from visual attention to gestural pointing), a major challenge this view faces is to show that there is more than a metaphorical sense in which the *formal* indexing involved in, e.g., anaphora resembles *physical* indexing, i.e., directing one's hearer's attention by pointing one's finger (or other body part) at a referent present in the discourse context. How does one get from the unconventionalized gestures which *accompany* speech to representations which *constitute* part of the linguistic capacity?

Recent work on sign-language may provide the missing link: Sign languages employ pointing gestures not only for deixis but also for anaphora. To refer to (possibly absent) individuals (persons, times, places or situations), signers first establish a *locus* for that individual in signing space using a nominal sign or a point. Subsequent reference to that same individual occurs by (literally) pointing back to that locus. Loci are "overt realization of indices" (Schlenker, 2011, p. 351). Although pronouns and anaphora are realized differently in signed and spoken languages, the same capacity is at play, i.e., pointing gestures in sign language are genuinely *linguistic* and even *grammatical*. Pronouns in SL obey similar binding constraints as their counterparts in other modalities (Meier & Lillo-Martin, 2013).

As Meier & Lillo-Martin (2013, p. 154) explain, "the study of pointing in signed languages gives us a window into how gestural elements become linguistic over time." Some sign-languages are young enough that we can observe their maturation from rudimentary homesigns to fully conventionalized languages. An example is Nicaraguan Sign Language (NSL), which appeared in the 1970s. Coppola & Senghas (2010) study the integration of indexical pointing into NSL. How the form is used by successive cohorts of signers "create[s] a record, like rings on a tree" (p. 548) of deixis' evolution. Results support the story in which co-speech pointing progressively looses its concrete spatial content and exophoric function, and gains abstract anaphoric and grammatical functions – a pattern also observed in diachronic crosslinguistic studies (Diessel, 1999).

Such research, whose results converge with neurolinguistic data suggesting that the processing of discourse reference relies on brain regions for spatial attention (Almor, Smith, Bonilha, Fridriksson, & Rorden, 2007), is a promising source of support for files' vertical extension. Nevertheless, challenges remain.

[28] Hurford (2003) proposes a similar three-stage model of the evolution of reference, starting with FINSTs and culminating in communicative reference.

Firstly, while there is evidence for the evolutionary step from communicative pointing to discourse reference, this is not the most crucial chapter in the story MFT tells. Although philosophers describe basic singular thoughts as 'demonstrative', there is a gap between the capacities required for *perceptual* tracking, and for understanding not only simple referential expressions like demonstratives, but even *pre-linguistic* forms of pointing found in all humans. *Declarative* pointing requires the metarepresentational capacity to grasp others' communicative intentions and make inferences based on a common ground, capacities which may be uniquely human (Tomasello, 2008). It is not enough to show that deictic gestures that accompany speech grow into linguistic elements; how we get from object-files to human pointing requires explanation.

Secondly, the fact that one representational category evolved from another only weakly suggests they are of the same psychological natural kind. This is the sense in which evolutionary or developmental evidence remains *indirect*: Kindhood is not defined purely historically, but requires shared psychological properties. In fact, putting too much weight on evolutionary considerations actually threatens MFT. Deixis is a basic linguistic form, which evolves into *many* other forms – a process known as "polygrammaticalization" (Diessel, 1999) As a result, developmental paths can be traced not only from deixis to singular discourse reference, but also to other forms which *prima facie* have nothing to do with the expression of *singular* thought. To give just two examples, in many languages *sentence connectives* and the *copula* have evolved from demonstratives (Diessel, 1999). Surely one would not wish to say that therefore, thoughts canonically expressed using such terms in the relevant languages involve the file-based indexing system, or are in any sense singular.

We conclude that the most ambitious version of MFT, according to which mental files are a wide-ranging psychological natural kind underlying all and only singular thinking, is unsupported by the available data. Defenders of MFT may have overestimated the similarities between different notions of 'file' used in philosophy and cognitive science. Nevertheless, critical examination of MFT opens up exciting avenues for further empirical research, especially concerning the relationship between our perceptual capacity to individuate and track basic individuals, and our higher-level capacities for singular thought. Mental files thus constitute a particularly promising field of interdisciplinary investigation, at the intersection of psychology, linguistics, and philosophy.

Bibliography

Almor, A., Smith, D. V., Bonilha, L., Fridriksson, J., & Rorden, C. (2007). What is in a name? Spatial brain circuits are used to track discourse references. *Neuroreport*, *18*(12), 1215-1219.

Alvarez, G. A., & Cavanagh, P. (2004). The capacity of visual short-term memory is set both by visual information load and by number of objects. *Psychological science*, *15*(2), 106-111.

Alvarez, G. A., & Cavanagh, P. (2005). Independent resources for attentional tracking in the left and right visual hemifields. *Psychological Science*, *16*(8), 637-643.

Alvarez, G. A., & Franconeri, S. L. (2007). How many objects can you track?: Evidence for a resource-limited attentive tracking mechanism. *Journal of Vision*, *7*(13), 14.

Anderson, D. E., & Awh, E. (2012). The plateau in mnemonic resolution across large set sizes indicates discrete resource limits in visual working memory. *Attention, Perception, & Psychophysics*, *74*(5), 891-910.

Anderson, J. R., & Reder, L. M. (1999). The fan effect: New results and new theories. *Journal of Experimental Psychology: General*, *128*(2), 186.

Austin, J. L. (1964). *Sense And Sensibilia; Reconstructed From The Manuscript Notes By G J Warnock*. Oxford University Press.

Bach, K. (1987). *Thought and Reference* (Vol. 40). Oxford University Press.

Bach, K. (2010). Getting a Thing into a Thought. In R. Jeshion (ed.), *New Essays on Singular Thought* (p. 39). Oxford University Press.

Bahrami, B. (2003). Object property encoding and change blindness in multiple object tracking. *Visual cognition*, *10*(8), 949-963.
Baillargeon, R. (1995). Physical reasoning in infancy. In M. S. Gazzaniga (ed.), *The Cognitive Neurosciences* (p. 181-204). Mit Press.
Ben-Shahar, O., Scholl, B. J., & Zucker, S. W. (2003). Where objects come from: Attention, segmentation, and textons. *Journal of Vision*, *3*(9), 474-474.
Blok, S., Newman, G., & Rips, L. J. (2005). Individuals and their concepts. *Categorization inside and outside the lab*, 127-149.
Bloom, P. (1998). Different structures for concepts of individuals, stuffs, and real kinds: One mama, more milk, and many mice. *Behavioral and Brain Sciences*, *21*(1), 66-67.
Bloom, P. (2002). *How children learn the meanings of words*. MIT press.
Bloom, P., & Gelman, S. A. (2008). Psychological essentialism in selecting the 14th Dalai Lama. *Trends in cognitive sciences*, *12*(7), 243-243.
Boghossian, P. A. (1994). The transparency of mental content. *Philosophical Perspectives*, *8*, 33-50.
Boyd, R. (1999). Homeostasis, species, and higher taxa. In R. A. Wilson (ed.), *Species: New Interdisciplinary Essays* (p. 141-85). Mit Press.
Boyer, P. (1998). If « tracking » is category-specific a « common structure » may be redundant. *Behavioral and Brain Sciences*, *21*(1), 67-68.
Burge, T. (2010). *Origins of Objectivity*. Oup Oxford.
Burke, L. (1952). On the tunnel effect. *Quarterly Journal of Experimental Psychology*, *4*(3), 121-138.
Burnston, D., & Cohen, J. (2012). Perception of Features and Perception of Objects. *Croatian Journal of Philosophy*, *12*(3), 283-314.
Cacchione, T., Indino, M., Fujita, K., Itakura, S., Matsuno, T., Schaub, S., & Amici, F. (2014). Universal ontology Attentive tracking of objects and substances across languages and over development. *International Journal of Behavioral Development*, *38*(6), 481-486.
Campbell, J. (1987). Is sense transparent? *Proceedings of the Aristotelian Society*, *88*(n/a), 273-292.
Campbell, J. (2011). Visual Attention and the Epistemic Role of Consciousness. In C. Mole, D. Smithies, & W. Wu (ed.), *Attention: Philosophical and Psychological Essays* (p. 323). Oxford University Press.
Carey, S. (2009). *The Origin of Concepts*. Oxford University Press.
Carey, S., & Spelke, E. (1996). Science and Core Knowledge. *Philosophy of Science*, *63*(4), 515 - 533.
Carey, S., & Xu, F. (2001). Infants' knowledge of objects: Beyond object files and object tracking. *Cognition*, *80*(1), 179-213.
Casati, R. (2004). Is the object concept formal? *Dialectica*, *58*(3), 383–394.
Chen, Z. (2012). Object-based attention: a tutorial review. *Attention, Perception, & Psychophysics*, *74*(5), 784-802.
Chen, Z., & Cave, K. R. (2008). Object-based attention with endogenous cuing and positional certainty. *Perception & Psychophysics*, *70*(8), 1435-1443.
Cheries, E. W., Mitroff, S. R., Wynn, K., & Scholl, B. J. (2008). Cohesion as a constraint on object persistence in infancy. *Developmental Science*, *11*(3), 427-432.
Coppola, M., & Senghas, A. (2010). The emergence of deixis in Nicaraguan signing. *Sign languages: A Cambridge language survey*, 543-569.
Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, *24*(1), 87-114.
Crane, T. (2011). The Singularity of Singular Thought. *Aristotelian Society Supplementary Volume*, *85*(1), 21-43.
Crawford, S. (2013). Object-Dependent Thought. In P. Harold (ed.), *Encyclopaedia of the Mind*. SAGE.
Cummins, R. (2010). *The World in the Head*. Oup Oxford.
Dickie, I. (2010). We are acquainted with Ordinary Things. In R. Jeshion (ed.), *New Essays on Singular Thought* (p. 213-45). Oxford University Press.
Diessel, H. (1999). *Demonstratives: form, function and grammaticalization* (Vol. 42). John Benjamins Publishing.
Driver, J., Davis, G., Russell, C., Turatto, M., & Freeman, E. (2001). Segmentation, attention and phenomenal visual objects. *Cognition*, *80*(1-2), 61-95.
Echeverri, S. (forthcoming). Object Files, Properties, and Perceptual Content. *Review of Philosophy and Psychology*, 1-25.
Egly, R., Driver, J., & Rafal, R. D. (1994). Shifting visual attention between objects and locations: evidence from normal and parietal lesion subjects. *Journal of Experimental Psychology: General*, *123*(2), 161.
Ereshefsky, M., & Reydon, T. A. C. (2015). Scientific kinds. *Philosophical Studies*, *172*(4), 969-986.
Feigenson, L., & Carey, S. (2005). On the limits of infants' quantification of small object arrays. *Cognition*, *97*(3), 295-313.
Flombaum, J. I., & Scholl, B. J. (2006). A temporal same-object advantage in the tunnel effect: Facilitated

change detection for persisting objects. *Journal of Experimental Psychology: Human Perception and Performance*, *32*(4), 840.
Flombaum, J. I., Scholl, B. J., & Santos, L. R. (2009). Spatiotemporal priority as a fundamental principle of object persistence. In B. M. Hood & L. Santos (ed.), *The Origins of Object Knowledge* (p. 135-164). Oxford University Press.
Fodor, J. A. (2008). *Lot 2: The Language of Thought Revisited*. Oxford University Press.
Franconeri, S. L., Alvarez, G. A., & Cavanagh, P. (2013). Flexible cognitive resources: competitive content maps for attention and memory. *Trends in Cognitive Sciences*, *17*(3), 134-141.
Franconeri, S. L., Jonathan, S. V., & Scimeca, J. M. (2010). Tracking multiple objects is limited only by object spacing, not by speed, time, or capacity. *Psychological Science*, *21*(7), 920-925.
Gallois, A. (2008). Identity over time. In *Stanford Encyclopedia of Philosophy*.
García-Carpintero, M. (2000). A presuppositional account of reference fixing. *Journal of Philosophy*, *97*(3), 109-147.
Gentner, D. (1982). Why nouns are learned before verbs: Linguistic relativity versus natural partitioning. *Center for the Study of Reading Technical Report; no. 257*.
Goodman, R. (forthcoming). Against the Mental Files Conception of Singular Thought. *Review of Philosophy and Psychology*, 1-25.
Hall, D. G. (2009). Proper names in early word learning: Rethinking a theoretical account of lexical development. *Mind and Language*, *24*(4), 404-432.
Hansen, C., & Rey, G. (forthcoming). Files and Singular Thoughts Without Objects or Acquaintance: The Prospects of Recanati's "Actualism. *Review of Philosophy and Psychology*, 1-16.
Hardman, K. O., & Cowan, N. (2014). Remembering complex objects in visual working memory: do capacity limits restrict objects or features?
Hawthorne, J., & Manley, D. (2012). *The Reference Book*. Oxford University Press.
Holcombe, A. O., & Chen, W.-Y. (2013). Splitting attention reduces temporal resolution from 7 Hz for tracking one object to< 3 Hz when tracking three. *Journal of Vision*, *13*(1), 12.
Holcombe, A. O., Chen, W.-Y., & Howe, P. D. (2014). Object tracking: Absence of long-range spatial interference supports resource theories. *Journal of vision*, *14*(6), 1.
Hollingworth, A., & Franconeri, S. L. (2009). The role of surface features in establishing object correspondence. *Cognition*, *113*, 150-166.
Hollingworth, A., & Rasmussen, I. P. (2010). Binding objects to locations: the relationship between object files and visual working memory. *Journal of Experimental Psychology: Human Perception and Performance*, *36*(3), 543.
Hommel, B. (1998). Event files: Evidence for automatic integration of stimulus-response episodes. *Visual Cognition*, *5*(1-2), 183-216.
Hommel, B. (2004). Event files: Feature binding in and across perception and action. *Trends in cognitive sciences*, *8*(11), 494-500.
Hood, B. M., & Bloom, P. (2008). Children prefer certain individuals over perfect duplicates. *Cognition*, *106*(1), 455-462.
Horowitz, T. S., & Cohen, M. A. (2010). Direction information in multiple object tracking is limited by a graded resource. *Attention, Perception, & Psychophysics*, *72*(7), 1765-1775.
Howard, C. J., & Holcombe, A. O. (2008). Tracking the changing features of multiple objects: Progressively poorer perceptual precision and progressively greater perceptual lag. *Vision research*, *48*(9), 1164-1180.
Huff, M., Papenmeier, F., & Zacks, J. M. (2012). Visual target detection is impaired at event boundaries. *Visual Cognition*, *20*(7), 848-864.
Huntley-Fenner, G., Carey, S., & Solimando, A. (2002). Objects are individuals but stuff doesn't count: Perceived rigidity and cohesiveness influence infants' representations of small groups of discrete entities. *Cognition*, *85*(3), 203-221.
Hurford, J. R. (2003). The neural basis of predicate-argument structure. *Behavioral and Brain Sciences*, *26*(3), 261-283.
Jeshion, R. (2009). The significance of names. *Mind and Language*, *24*(4), 370-403.
Jeshion, R. (2010). Singular thought: acquaintance, semantic instrumentalism, and cognitivism. In R. Jeshion (ed.), *New Essays on Singular Thought* (p. 105-141). Oxford University Press.
Jeshion, R. (2014). Two Dogmas of Russellianism. In *Empty Representations*. Oxford: Oxford University Press.
Jordan, K. E., Clark, K., & Mitroff, S. R. (2010). See an object, hear an object file: Object correspondence transcends sensory modality. *Visual Cognition*, *18*(4), 492-503.
Kahneman, D., & Treisman, A. (1984). Changing views of attention and automaticity. *Varieties of attention*, *1*, 29-61.
Kahneman, D., Treisman, A., & Gibbs, B. J. (1992). The reviewing of object files: Object-specific integration of information. *Cognitive psychology*, *24*(2), 175-219.

Kamp, H. (1984). Context, Thought and Communication. *Proceedings of the Aristotelian Society*, *85*(n/a), 239 - 261.
Kim, J. (1977). Perception and reference without causality. *Journal of Philosophy*, *74*(October), 606-620.
Kripke, S. A. (1980). *Naming and Necessity* (Vol. 217). Harvard University Press.
Krøjgaard, P. (2000). Object individuation in 10-month-old infants: Do significant objects make a difference? *Cognitive Development*, *15*(2), 169-184.
Lawlor, K. (2001). *New Thoughts About Old Things: Cognitive Policies as the Ground of Singular Concepts*. Garland Pub.
Lewis, D. (1984). Putnam's paradox. *Australasian Journal of Philosophy*, *62*(3), 221 – 236.
Liverence, B. M., & Scholl, B. J. (2011). Selective Attention Warps Spatial Representation Parallel but Opposing Effects on Attended Versus Inhibited Objects. *Psychological science*, *22*(12), 1600-1608.
Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, *390*(6657), 279-281.
Luck, S. J., & Vogel, E. K. (2013). Visual working memory capacity: from psychophysics and neurobiology to individual differences. *Trends in Cognitive Sciences*, *17*(8), 391-400.
Luo, Y., & Baillargeon, R. (2005). When the ordinary seems unexpected: evidence for incremental physical knowledge in young infants. *Cognition*, *95*(3), 297-328.
Magliano, J. P., & Zacks, J. M. (2011). The Impact of Continuity Editing in Narrative Film on Event Segmentation. *Cognitive Science*, *35*(8), 1489-1517.
Marcus, G. F. (2001). *The Algebraic Mind*. The Mit Press.
Margolis, E. (1998). How to acquire a concept. *Mind and Language*, *13*(3), 347-369.
Meier, R. P., & Lillo-Martin, D. (2013). The points of language.
Millikan, R. G. (1998). A common structure for concepts of individuals, stuffs, and real kinds: More Mama, more milk, and more mouse. *Behavioral and Brain Sciences*, *21*(1), 55-65.
Millikan, R. G. (2000). *On Clear and Confused Ideas: An Essay About Substance Concepts* (Vol. 65). Cambridge University Press.
Mitroff, S. R., & Alvarez, G. A. (2007). Space and time, not surface features, guide object persistence. *Psychonomic Bulletin & Review*, *14*(6), 1199-1204.
Mitroff, S. R., Scholl, B. J., & Wynn, K. (2004). Divide and conquer how object files adapt when a persisting object splits into two. *Psychological Science*, *15*(6), 420-425.
Mitroff, S. R., Scholl, B. J., & Wynn, K. (2005). The relationship between object Files and conscious perception. *Cognition*, *96*(1), 67-92.
Montague, M. (2011). The phenomenology of particularity. In T. Bayne & M. Montague (ed.), *Cognitive Phenomenology* (p. 121-140). Oxford University Press.
Moore, C. M., Stephens, T., & Hein, E. (2010). Features, as well as space and time, guide object persistence. *Psychonomic Bulletin & Review*, *17*(5), 731-736.
Murez, M., & Smortchkova, J. (2014). Singular Thought: Object□Files, Person□Files, and the Sortal PERSON. *Topics in cognitive science*, *6*(4), 632-646.
Newtson, D., & Engquist, G. (1976). The perceptual organization of ongoing behavior. *Journal of Experimental Social Psychology*, *12*(5), 436-450.
Noles, N. S., Scholl, B. J., & Mitroff, S. R. (2005). The persistence of object file representations. *Perception & Psychophysics*, *67*(2), 324-334.
Norman, L. J., Heywood, C. A., & Kentridge, R. W. (2013). Object-based attention without awareness. *Psychological science*, *24*(6), 836-843.
Odic, D., Roth, O., & Flombaum, J. I. (2012). The relationship between apparent motion and object files. *Visual Cognition*, *20*(9), 1052-1081.
Oksama, L., & Hyönä, J. (2004). Is multiple object tracking carried out automatically by an early vision mechanism independent of higher□order cognition? An individual difference approach. *Visual cognition*, *11*(5), 631-671.
Oksama, L., & Hyönä, J. (2008). Dynamic binding of identity and location information: A serial model of multiple identity tracking. *Cognitive psychology*, *56*(4), 237-283.
Orlandi, N. (2014). *The Innocent Eye: Why Vision is Not a Cognitive Process*. Oup Usa.
Peacocke, C. (2001). Does perception have a nonconceptual content? *Journal of Philosophy*, *98*(5), 239-264.
Perry, J. (1980). A Problem About Continued Belief. *Pacific Philosophical Quarterly*, *61*(4), 317.
Perry, J. (1993). *The Problem of the Essential Indexical: And Other Essays* (Vol. 44). Oxford University Press.
Pinto, Y., Howe, P. D., Cohen, M. A., & Horowitz, T. S. (2010). The more often you see an object, the easier it becomes to track it. *Journal of vision*, *10*(10), 4.
Prinz, J. J. (2005). The return of concept empiricism. In H. Cohen & C. Leferbvre (ed.), *Categorization and Cognitive Science*. Elsevier.
Prosser, S. (2005). Cognitive dynamics and indexicals. *Mind and Language*, *20*(4), 369–391.

Pylyshyn, Z. (2004). Some puzzling findings in multiple object tracking: I. Tracking without keeping track of object identities. *Visual cognition*, *11*(7), 801-822.
Pylyshyn, Z. W. (2007). *Things and Places: How the Mind Connects with the World*. The Mit Press.
Radvansky, G. A. (1999). The fan effect: a tale of two theories.
Radvansky, G. A., & Zacks, J. M. (2011). Event perception. *Wiley Interdisciplinary Reviews: Cognitive Science*, *2*(6), 608-620.
Recanati, F. (1993). *Direct Reference: From Language to Thought*. Blackwell.
Recanati, F. (2005). Deixis and Anaphora. In Z. G. Szabó (ed.), *Semantics Versus Pragmatics* (p. 286-316). Oxford University Press.
Recanati, F. (2010). Singular Thought: In Defense of Acquaintance. In R. Jeshion (ed.), *New Essays on Singular Thought*. Oup Oxford.
Recanati, F. (2012). *Mental Files*. Oxford University Press.
Recanati, F. (2013). Mental Files: Replies to my Critics. *Disputatio*, *5*(36).
Rensink, R. A., & Enns, J. T. (1998). Early completion of occluded objects. *Vision research*, *38*(15), 2489-2505.
Rips, L. (2011). *Lines of Thought: Central Concepts in Cognitive Psychology*. Oup Usa.
Rips, L. J., & Hespos, S. J. (2015). Divisions of the Physical World: Concepts of Objects and Substances.
Rugani, R., Fontanari, L., Simoni, E., Regolin, L., & Vallortigara, G. (2009). Arithmetic in newborn chicks. *Proceedings of the Royal Society of London B: Biological Sciences*, *276*(1666), 2451-2460.
Sainsbury, M. (2005). *Reference Without Referents*. Clarendon Press.
Sawyer, S. (2012). Cognitivism: A New Theory of Singular Thought? *Mind and Language*, *27*(3), 264-283.
Saxe, R., Tenenbaum, J. B., & Carey, S. (2005). Secret agents inferences about hidden causes by 10-and 12-month-old infants. *Psychological Science*, *16*(12), 995-1001.
Schiffer, S. (1978). The basis of reference. *Erkenntnis*, *13*(1), 171-206.
Schiffer, S. (1990). The Mode-of-Presentation Problem. In C. A. A. J. Owens (ed.), *Propositional Attitudes: The Role of Content in Logic, Language, and Mind* (p. 249-268). CSLI.
Schiffer, S. (1996). Book review. *Linguistics and Philosophy*, *19*(1), 91-102.
Schlenker, P. (2011). Donkey anaphora: the view from sign language (ASL and LSF). *Linguistics and Philosophy*, *34*(4), 341-395.
Scholl, B. (2001). Objects and attention: the state of the art. *Cognition*, *80*(1-2), 1-46.
Scholl, B. J. (2007). Object persistence in philosophy and psychology. *Mind and Language*, *22*(5), 563–591.
Scholl, B. J. (2009). What have we learned about attention from multiple object tracking (and vice versa). *Computation, cognition, and Pylyshyn*, 49-78.
Scholl, B. J., & Leslie, A. M. (1999). Explaining the infant's object concept: Beyond the perception/cognition dichotomy. *What is cognitive science*, 26-73.
Scholl, B. J., & Xu, Y. (2001). The magical number 4 in vision. *Behavioral and Brain Sciences*, *24*(1), 145-146.
Schroeter, L. (2007). Illusion of transparency. *Australasian Journal of Philosophy*, *85*(4), 597 – 618.
Scimeca, J. M., & Franconeri, S. L. (2015). Selecting and tracking multiple objects. *Wiley Interdisciplinary Reviews: Cognitive Science*, *6*(2), 109-118.
Searle, J. R. (1983). *Intentionality: An Essay in the Philosophy of Mind*. Cambridge University Press.
Shipley, T. F., Shipley, T. F., & Zachs, J. M. (2008). An invitation to an event. *Understanding Events: From Perception to Action: From Perception to Action*, 3-30.
Spelke, E. S. (1990). Principles of object perception. *Cognitive science*, *14*(1), 29-56.
Spelke, E. S., & Kinzler, K. D. (2007). Core knowledge. *Developmental science*, *10*(1), 89-96.
Strickland, B., & Keil, F. (2011). Event completion: Event based inferences distort memory in a matter of seconds. *Cognition*, *121*(3), 409-415.
Swallow, K. M., Zacks, J. M., & Abrams, R. A. (2009). Event boundaries in perception affect memory encoding and updating. *Journal of Experimental Psychology: General*, *138*(2), 236.
Taylor, K. A. (2003). *Reference and the Rational Mind*. CSLI Publications.
Tomasello, M. (2008). *Origins of Human Communication*. MIT Press.
Tripathy, S. P., Ogmen, H., & Narasimhan, S. (2011). Multiple-object tracking: A serial attentional process. In C. Mole, D. Smithies, & W. Wu (ed.), *Attention: Philosophical and Psychological Essays* (p. 117-144). Oxford University Press.
Tsompanidis, V. (2015). Mental Files and Times. *Topoi*, *34*(1), 233-240.
Van Marle, K., & Scholl, B. J. (2003). Attentive tracking of objects ys. substances. *Psychological Science*, *14*.
Vickery, T. J., & Chun, M. M. (2010). Object-based warping an illusory distortion of space within objects. *Psychological science*.
Vogel, E. K., Woodman, G. F., & Luck, S. J. (2001). Storage of features, conjunctions, and objects in visual working memory. *Journal of Experimental Psychology: Human Perception and Performance*, *27*(1), 92.
Wood, J. N. (2007). Visual working memory for observed actions. *Journal of Experimental Psychology: General*, *136*(4), 639.

Xu, F. (2007). Sortal concepts, object individuation, and language. *Trends in Cognitive Sciences*, *11*(9), 400-406.
Xu, F., & Carey, S. (1996). Infants' metaphysics: The case of numerical identity. *Cognitive psychology*, *30*(2), 111-153.
Xu, Y., & Chun, M. M. (2009). Selecting and perceiving multiple visual objects. *Trends in Cognitive Sciences*, *13*(4), 167-174.
Zacks, J. M., & Swallow, K. M. (2007). Event segmentation. *Current Directions in Psychological Science*, *16*(2), 80-84.