

# Perceptual modes of presentation as object files

\*Forthcoming in *Erkenntnis*. Please cite published version when possible.

Gabriel Siegel

Washington University in St. Louis

## Abstract

Some have defended a Fregean view of perceptual content. On this view, the constituents of perceptual contents are Fregean *modes of presentation* (MOPs). In this paper, I propose that perceptual MOPs are best understood in terms of *object files*. Object files are episodic representations that store perceptual information about objects. This information is updated when sensory conditions change. On the proposed view, when a subject perceptually represents some object *a* under two distinct MOPs, then the subject initiates two object files that both refer to *a*. My defense of this view appeals to its satisfaction of four constraints that I argue theories of perceptual MOPs should satisfy. Furthermore, I show that some existent accounts of perceptual MOPs fail to satisfy them. The defended constraints also indicate what is unique about perceptual, as opposed to linguistic or cognitive, MOPs.

## Introduction

Some have defended a Fregean view of perceptual content. On this view, the constituents of perceptual contents are Fregean *modes of presentation* (MOPs). In this paper, I propose that perceptual MOPs are best understood in terms of *object files*. Object files are a construct currently used in cognitive science and first developed by Kahneman, Treisman and Gibbs (1992). These files are episodic representations that store information about objects in one's perceptual field, and this information is updated when sensory conditions change. On the proposed view, token perceptual MOPs are object file representations in perceptual systems. Perceptual MOPs are understood as abstractions from those token representation instances. When a subject perceptually represents some object *a* under two distinct MOPs, then the subject initiates two object files that both refer to *a*. To defend this view, I argue for a set of constraints on any theory of perceptual MOPs. I show that the object files view satisfies all the specified constraints.

The following discussion gives an account of perceptual MOPs of *objects*. I ultimately leave it as an open-ended question whether there are MOPs of perceptual attributes and events.<sup>1</sup> Furthermore, on the assumption that there are MOPs of attributes and events in perception, I leave it as open-ended how the proposed account might be applied to them. MOPs of attributes and events raise another set of questions.<sup>2</sup> I do not address these questions in the present paper.

In section 1.1, I outline the distinctive epistemic situation that agents are in when they perceive an object under different MOPs. Then, in section 1.2, I distinguish perceptual MOPs from what Crane (2003) calls *intentional modes*. As I will argue, accounts that conflate intentional modes with MOPs are not feasible. In section 1.3, I argue that perceptual MOPs are not individuated by the representation of distinct properties attributed to their referent. For example, a representation of an object *a* with the attributed property F does not necessarily constitute a distinct MOP from a representation of *a* with the attributed property G. In section 1.4, I argue that a theory of perceptual MOPs should elucidate how the content of perceptual experiences enables singular thinking. These four subsections provide a framework for analyzing the plausibility of theories of perceptual MOPs. Furthermore, the framework indicates what is unique about perceptual, as opposed to linguistic or cognitive, MOPs.

After providing this framework, in section 2, I give an object files account of perceptual MOPs and I show how the account meets all the specified constraints. Finally, section 3 looks at some existent views on what perceptual MOPs are, i.e., Chalmers (2004) and Schellenberg (2018). In this section, I show that these views fail to meet some of the constraints specified in section 1.

## **1. Constraints on theories of perceptual MOPs**

### **1.1 Frege's criterion and perceptual Frege cases**

It is not completely clear what Frege meant by MOPs. However, what is uncontroversial is his criterion for a *difference* in MOP. According to Frege, two linguistic expressions can have the same reference but different *senses* (or MOPs). For example, the two expressions 'morning star' ('MS') and 'evening star' ('ES') both refer to Venus, but each presents Venus to an agent differently. If the agent does not know that 'MS' co-refers with 'ES,' then they might accept (a)

---

<sup>1</sup> For discussions of perceptual MOPs of attributes and property-instances, see Chalmers (2004), Thompson (2009), Burge (2010), Schellenberg (2018), and Rescorla (2020). For a discussion of perceptual MOPs of events, see Schellenberg (2018).

<sup>2</sup> For example, how perceptual contents can have distinct objectual constituents but correspond with the same phenomenal character, see Thompson (2009).

“MS is a body illuminated by the sun,” while rejecting (b) “ES is a body illuminated by the sun.” Since it is possible for a rational agent to accept (a), while rejecting (b), given that the only difference is that ‘MS’ is replaced by ‘ES,’ these expressions meet Frege’s criterion for having different MOPs (1892: 199-202).

The notion of a linguistic MOP explains why the agent is in a special kind of epistemic situation. This situation has three important characteristics. First, thinkers can believe that  $MS = MS$  without believing that  $MS = ES$ . This allows rational agents to attribute a property to MS without attributing that same property to ES. Second, the agent is in a position to learn something new when they are informed that MS refers to the same object as ES. Here the identity proposition expressed by ‘ $MS = ES$ ’ is *informative*. Third, the agent, in virtue of speaking about Venus via two distinct MOPs, as well as being unaware that they are co-referential, is in a position to make inconsistent and contradictory claims about the same object. They can make these claims without losing their rational standing. These three epistemic features are understood in virtue of the expressions’ propositional contents involving distinct, but co-referential, MOPs. If the agent’s mental state did not have this Fregean content, then they would not be in this kind of epistemic situation. Hence, co-referential MOPs explain the special epistemic circumstance faced by agents in these types of cases.

Frege’s criterion should also apply to perceptual MOPs. Consider two distinct perceptual MOPs, e.g., *a* and *b*, that refer to the same object. If an agent’s perceptual system represents both MOPs and is unaware that they co-refer, then it is possible for the system to attribute a feature to *a* but not attribute that same feature to *b*. Furthermore, it is possible for the system to instantiate inconsistent attribution representations for the same object. By *inconsistent*, I mean that the attributions of these features to the same object cannot be true together. This possibility is the perceptual analog to Frege cases at the level of thought. For example, one’s perceptual system might represent, in perceptual consciousness, that *a* instantiates some feature F, and *b* instantiates G, where F and G are inconsistent and  $a = b$ . I assume that the perceptual system doesn’t need to represent these properties *as contradictory* in order for a Frege case to arise. Inconsistent perceptual attributions can be re-expressed, at the level of thought, as a commitment to a

contradiction.<sup>3</sup> In this way, the perceptual attribution of inconsistent features to the same object suffices for a Frege case.

If representing the same object in two different ways can lead to these kinds of mistakes, then those two different ways satisfy Frege's criterion for distinct MOPs (concrete examples will be provided shortly). Like the Venus case above, perceptual MOPs need to explain why the perceiver is in this special kind of epistemic position.

To understand how perceptual Frege cases can arise requires understanding what perceptual systems can do. We want to understand the conditions under which a perceptual system might represent a Frege-style inconsistency in perceptual consciousness. The scope of these conditions will depend on *perceptual capacities*.<sup>4</sup> We can broadly think of perceptual capacities as capacities to be sensitive to objects and properties in the environment across varying conditions. These perceptual capacities respond to those object and properties in such a way that grounds cognitive and motor activities (O'Callaghan 2019: 56). Many perceptual capacities represent objects and properties in perceptual consciousness.

Perceptual MOPs are understood as ways in which some object is presented such that a Frege case can arise in perception. In this vein, the conditions under which a perceptual Frege case can arise constrains our understanding of what perceptual MOPs are. Understanding the conditions under which perceptual Frege cases arise requires understanding perceptual capacities. In particular, how a perceptual system is capable of attributing inconsistent properties to the same object. Thus, the set of objectual appearances that qualify as perceptual MOPs are limited by the range of perceptual capacities.

Consider an example. You perceive some object *a* today and perceptually attribute the quality F to it. Then, tomorrow, you perceive *a* again, mistake it for a different object, and attribute -F to it. You are now in a position to have the thought expressed by '*Fa* and *-Fa*.' This is not a *perceptual* Frege case because it would involve retrieving information stored in long-term memory, a cognitive capacity. To be confused about whether two objects are the same, when they are not represented in perceptual experience simultaneously, requires a cognitive capacity to represent objects that one is not currently perceptually related to.<sup>5</sup> By contrast, a perceptual

---

<sup>3</sup> For example, consider the inconsistent perceptual attribution expressed by '*a* is to the left of the streetlight' and '*b* is to the right of the streetlight' where *a* = *b*. This could be re-expressed in thought as '*a* is to the left of the streetlight' and '*b* is *not* to the left of the streetlight.'

<sup>4</sup> A perceptual capacity has been understood in different ways. For example, as a capacity to discriminate and single out particulars (Schellenberg 2018) or to detect and differentiate particulars and features (O'Callaghan 2019).

<sup>5</sup> I am assuming that the identity confusion happens after representations stored in perceptual short-term memory.

system's capacity to represent is *stimulus-dependent*. Since this case involves an extra-perceptual cognitive capacity, this is merely a Frege case at the level of thought. The different ways in which *a* is presented to you, today and tomorrow, are thus only cognitive MOPs for *a*, not perceptual ones. This is the case because these different ways of presenting *a* cannot lead to attributing inconsistent properties to *a* in a way that one's perceptual system is capable of.

For the remainder of this subsection, I provide a characterization of perceptual Frege cases. In order for a Frege case to arise in perception, there are some broad conditions that need to be met. First, the subject represents some perceptual thing, or things, as having some feature (or features). This assumes that perception often has an attributional structure (Burge 2010). According to this assumption, much of our perceptual experiences constitutively involve the attribution of properties to objects.<sup>6</sup> Second, the subject is in a position such that they might represent there being two (or more) objects when there is really only one (see Millar 2015: 372). Third, the subject is in a position to make inconsistent attribution representations for the same object. In other words, the subject is in a position to make the perceptual analog to Frege type identity mistakes. This is the unique epistemic situation specified above. To better understand what perceptual Frege cases are, I'll consider some examples.

Consider that you are intoxicated and seeing double. Some object *a* appears to you twice (as *a* and *b*). You represent a spatial relation between *a* and *b*.<sup>7</sup> When a perceptual system represents a spatial relation between *a* and *b*, we can say that the perceptual system treats *a* and *b* as distinct (i.e., non-identical) objects. Given the spatial relation between the two objects, your perceptual experience might represent *a* and *b* as having inconsistent spatial features. For example, '*a* is to the right of the streetlight' and '*b* is to the left of the streetlight,' when *a* = *b*. Since objects cannot be bilocated, these attribution representations are inconsistent. The notion of perceptual MOPs helps explain why the perceptual system attributes the feature 'to the right of the streetlight' to *a* but not to *b*.

John Perry (2001) provides an analogous case where there is a dog with a long torso wrapped around a post. A perceiver might falsely represent the one dog as two dogs. What appear to be two dogs are really two parts of the same dog. Thus, one's perceptual experience might represent inconsistent spatial properties as being attributed to the same object.

---

<sup>6</sup> As further motivation for this assumption, many cases of feature binding, e.g., Treisman (1998), are plausibly construed as attributional.

<sup>7</sup> This is an ideal case of seeing double, where the perceptual system represents two objects in two distinct spatial locations.

These cases involve non-veridical attributions of spatial properties to an object in vision. Other kinds of perceptual Frege cases might involve the way in which we attribute spatial properties across different sensory modalities. For example, consider a visual representation of a tennis ball to the left of a pole. In addition, consider, due to an echo effect, a simultaneous auditory representation of the same ball as to the right of the pole.<sup>8</sup> Like the cases above, this crossmodal case involves the attribution of inconsistent spatial properties to the same object.<sup>9</sup>

Any account of perceptual MOPs should explain why the cases illustrated above are Frege cases. It should explain that the subject is in an epistemic position such that they can make perceptual identity mistakes because they are representing the same object via multiple distinct MOPs.

## 1.2 MOPs and intentional modes

Why can't distinct perceptual MOPs be understood as perceiving objects via distinct sensory modalities? For example, why doesn't hearing *a* and seeing *a* constitute two different MOPs for *a*? These different ways of perceiving are a subset of what I will call, following Crane (2003), different perceptual *intentional modes*. Once we see that a principled distinction is to be made between perceptual intentional modes and perceptual MOPs, it will be apparent that MOPs cannot be identified with intentional modes.

Crane makes a distinction between intentional content and intentional modes. Intentional content is what a particular mental state is about (2003: 39). For our purposes, the intentional content of perceptual states specifies what is represented to the subject during perceptual experiences. In other words, the content has a propositional structure and represents the world as being a particular way. Intentional modes specify the way in which a subject relates to some content. This might be a belief relation, a desiring relation, or some kind of perceptual relation (2003: 39). Chalmers makes the same distinction as one between content and the way in which that content is represented. He calls the latter *manners of representation* (2004: 155).

For present purposes, I will assume, following Chalmers (2004), that a full specification of the phenomenology of perceptual episodes involves specifying both the content and the intentional mode(s). Theorists who reject this assumption will not find the view that perceptual MOPs are

---

<sup>8</sup> I assume that spatial attributions made in both vision and audition can involve the same object.

<sup>9</sup> I leave it as an open empirical question whether there are perceptual Frege cases beyond the attribution of inconsistent spatial properties.

intentional modes plausible to begin with.<sup>10</sup> As Chalmers suggests, perceptual intentional modes can be more or less fine-grained (2004: 155). For example, a visual perception with a certain representational content *C* will have a different phenomenology from an auditory perception with content *C*. This divergence in phenomenology can be accounted for by an auditory and visual intentional mode, both with the content *C*, respectively.<sup>11</sup>

In language, the statements ‘*MS is F*’ and ‘*ES is F*’ represent distinct information with distinct cognitive ‘values’ or ‘significances’ (Frege 1892). This is the case even though ‘*MS*’ and ‘*ES*’ both refer to Venus. One reason is that *MS = ES* is *informative*. This is the case since, unlike  $a = a$ , it is possible for a subject to be unaware that *MS = ES* and thus learn something new when they are informed that *MS = ES*. Since *MS = ES* is informative, we can assume that the information represented to subjects by ‘*MS is F*’ and ‘*ES is F*’ is different, even though the contents ascribe the same property to the same object. In this vein, the informativity of identity propositions with distinct MOPs is an important characteristic of MOPs.

As in the case of thought, it is possible for a perceptual system to become responsive to the fact that two MOPs of the same object are co-referential. Consider again Perry’s dog around the pole case where a spatial relation is represented between what appears to be two dogs *a* and *b*. If the perceptual system is provided with new information, it can become responsive to the fact that *a* is co-referential with *b*. This kind of transition is the perceptual analog to ‘learning something new’ at the level of thought. For example, if the dog’s head and tail suddenly become visible, in the appropriate way, the perceptual system might transition from representing the two dogs to representing only one. In this way, perceptual MOPs like *a* and *b* represent distinct information because a perceptual system can be unaware, and then ‘learn,’ that  $a = b$ .<sup>12</sup> Here, ‘learning’ plausibly means either that the perceptual system: 1) transitions from representing two objects to representing one or 2) continues to represent two distinct objects but is sensitive to the fact that they are co-referential.

I don’t commit, in the present paper, to what the transition in (2) entails specifically. However, O’Callaghan’s (forthcoming) discussion of crossmodal identification provides a

---

<sup>10</sup> For arguments against this view see O’Callaghan (2014) and Bourget (2017).

<sup>11</sup> Speaks, on the other hand, has appealed to intermodal cases in order to argue against dividing up perceptual intentional modes in this way (see 2015: 179-185).

<sup>12</sup> In a recent discussion, Rescorla (2020) discusses cases of perceptual co-reference from the perspective of Bayesian perceptual psychology. For example, both haptic and visual representations, i.e., Bayesian ‘hypotheses,’ of the size of a distal object can be distinct representations but co-referential for the perceptual system. This is an example, Rescorla suggests, of co-reference of *features* of perceptual objects.

plausible characterization of the transition under consideration. This is the transition involved in a perceptual system becoming sensitive to the fact that distinct MOPs are co-referential. Roughly, this entails that the perceptual system detects that  $a = b$  and responds differently than it would if it did not detect that  $a = b$ . This does not entail that the perceptual system, in ‘learning something new,’ must represent that  $a = b$ . O’Callaghan’s characterization implicates the notion of ‘differential sensitivity,’ where the system responds differently to instances in which intermodal objects are co-referential and instances in which they are not.

To illustrate, *attentional spreading* is an example where the perceptual system responds differently to there being one object as opposed to multiple. Attentional spreading is the enhanced perceptual processing of features of an object that is attended to. The enhancement is a function of attention ‘spreading’ throughout the attended object. However, this ‘spreading’ does not extend to distinct surrounding objects. For this reason, the enhanced processing is not present for features of distinct objects surrounding the attended object. This is the case even when an object exhibiting attentional spreading is later broken up into multiple objects. According to O’Callaghan (forthcoming), this demonstrates perceptual sensitivity to whether there are one or multiple objects.<sup>13</sup> To experimentally demonstrate transition (2), it is possible for attentional spreading paradigms to indicate sensitivity to co-reference (i.e., there being only one object) while a conscious representation of two objects persists (see footnote).<sup>14</sup>

I’ve shown that representing distinct MOPs entails a difference in what information is represented to perceiving subjects. This is the case since a perceptual system can become sensitive to the fact that two MOPs are co-referential. On the other hand, varying intentional modes do not imply a difference in what information is represented to subjects. Consider that a perceptual system produces a conscious auditory representation of the spatial location of  $a$  and then visually represents  $a$  as being in the exact same region of space. The latter perception, given the intentional

---

<sup>13</sup> In Richard et al.’s (2008) experiment, the stimuli were long solid rods with six ‘bites’ (or ‘carvings’) taken out. These bites were either bars or half-circles. Sometimes the stimulus was a single object and sometimes the stimulus was broken up into three objects. The task was to identify the shape of the middle two bites (either bars or half-circles). In the single object condition, reaction times were faster when the middle shape bites (the targets) matched the surrounding other four bites. Alternatively, reaction times were slower when the target was different from the surrounding bites. However, when the object was broken up into three objects, these effects went away. According to O’Callaghan (forthcoming), this is a case that demonstrates differential sensitivity to co-reference. The faster reaction times when there are matching bites, in the single object condition, but not in the multiple objects condition, suggests differential sensitivity to there being one object as opposed to multiple.

<sup>14</sup> Richard et al.’s (2008) design could be turned into a perceptual Frege case. In this hypothetical adaptation, the single object stimulus would be separated by a vertical post in the exact middle of the object. Like the dog case, this object could either be represented as two objects or one. Consider that the subject has a conscious perceptual representation of two distinct objects but exhibits reaction times that are slower when the surrounding bites don’t match the targets (and faster when they do match). These reaction times should be comparable to the results noted above in the single object condition as opposed to the multiple objects condition. In this way, the presence or absence of the reaction time effects would indicate whether the perceptual system is sensitive to the fact that the two objects are co-referential. This sensitivity to co-reference can be indicated while subjects report a conscious perceptual representation of two distinct objects.



modes framework, does not provide new information regarding the numerical identity of  $a$  or what spatial features are attributed to  $a$ .<sup>15</sup> This is the case because, in both cases, the perceptual content is assumed to be the same. This is unlike when a perceptual system is provided with new information regarding the co-reference of two MOPs.

In short, if we adopt Crane and Chalmers's intentional modes framework, MOPs help constitute the *content* of perceptual episodes, not the relation subjects bear to contents. For example, the sensory modalities in virtue of which perceptual information is represented is orthogonal, given the intentional modes framework, to the question of what information is represented. As I've shown above, perceptual MOPs help constitute the latter, not the former. For this reason, perceptual MOPs should not be conflated with perceptual intentional modes.

### 1.3 Perceptual MOPs and representations of attributes

In the next two sections, I motivate two related claims. First, that perceptual MOPs are not individuated by the representation of distinct properties attributed to their referent. Second, that perceptual MOPs enable singular (or '*de re*') thoughts. The first claim is that different MOPs are not individuated in virtue of distinct representations, where the only difference between these representations is that different attributes are ascribed to the same referent. For example, a representation of an object  $a$  with the attributed property  $F$  does not necessarily constitute a distinct MOP from a representation of  $a$  with the attributed property  $G$ .

To illustrate this idea, consider cases where the representation of distinct features does not intuitively place subjects into the epistemic circumstance outlined in section 1.1. For example, consider very slight alterations in illumination features, common differences in spatial features (e.g., when an object moves slightly across one's visual field) and alterations in spatial perspectival features as an agent moves around an object. These are differences in the representation of features attributed to some object (e.g.,  $a$ ). However, it is counterintuitive to think that these differences automatically make subjects prone to make Frege-style identity mistakes. For example, they do not intuitively create the possibility for confusion about whether  $a$  at  $t^1$  is the same object encountered at  $t^2$ . In addition, if each alteration of the representation of features attributed to  $a$

---

<sup>15</sup> We might think that they learn something new *phenomenologically*, i.e., about what it's like to undergo visual experiences of ' $a$  is  $F$ ' rather than auditory ones. However, this kind of epistemic gain is not what is under consideration.

consequences in a new MOP for *a*, then perceiving subjects are constantly presented with novel MOPs for *a*. Under this view, any object can in principle have an infinite number of MOPs.

It is intuitive to think that this is implausible, i.e., that what individuates an object's MOP is not *the features* associated with the object, but *how those features are organized* in one's mental economy. In other words, what explains the possibility of Frege cases is not merely the representation of different features. Rather, distinct MOPs result from how attributional information is organized, not the attributional information itself. This insight has been carefully developed in the mental files (MFs) literature. For example, Recanati (2012) suggests that a *de re* cognitive MOP is not the collection of predicates an agent believes some referent satisfies, but rather a MF itself (2012: 40-41). MFs are designed to exploit *acquaintance relations*, i.e., "information-providing relations," and store the information acquired by them (2012: 36-37, 60). It is these acquaintance relations that individuate MFs, not the information that they store. (In section 2, I discuss the MF framework in greater detail.)

A couple responses are open to the theorist who denies this constraint. First, it might be held that only certain represented properties, or certain represented properties in certain contexts, constitute a new MOP for *a*. While a live option, such theorists would need a principled way to distinguish which represented properties do, and do not, constitute a new MOP for objects. Second, it might be argued that, even slight differences in illumination, or differences in spatial properties, can lead subjects to make Frege-style identity mistakes in some possible scenario. In other words, for any difference in attributional representations there is some possible world in which that difference leads to a Frege case for perception. While possible, it is implausible that, in normal contexts in this world, we become confused about co-reference solely on the basis of such minor representational changes.

If the above is right, then perceptual MOPs are not individuated in virtue of the representation of distinct properties attributed to the same object. An account of perceptual MOPs must specify how each MOP contains information about numerical identity. That is, the MOP contains some way to individuate itself from other co-referential MOPs. In this section, I've argued that they do this in such a way that does not rely on what features the subject attributes to the object.

## 1.4 Enabling *de re* thought

In this section, I argue that a theory of perceptual MOPs should elucidate how the representational contents of perceptual experiences enables *de re*<sup>16</sup> thinking. To have a *de re* or singular thought is to think about something in a way that is unmediated by descriptions (or any way of conceptualizing the object). To think about an object *de re* is to think about the object in such a way that the subject predicates qualities of *it*. The thoughts that accompany perceptual experiences are widely viewed as paradigmatic of *de re* thinking. For example, Kaplan (1989) argued that perception of objects grounds demonstrative thinking (a subset of *de re* thoughts). When you perceive an object, you can have a demonstrative thought of the kind ‘*that* is to the left of the streetlight.’<sup>17</sup>

In order for the object constituent of thought contents to be non-descriptive, it is natural to think that the object constituent of perceptual contents, which enables the *de re* thought contents, should also be non-descriptive. If this were not the case, it would be unclear how the *de re* thought gets its non-descriptive character. Hence, perceptual MOPs of objects are plausibly non-descriptive (i.e., *de re*) content constituents. This means that their mechanisms for fixing reference are not satisfactorily, i.e., the MOPs don’t refer to objects in virtue of satisfying a description. Furthermore, *de re* MOPs are not individuated in virtue of distinct descriptions, but something else (e.g., acquaintance relations).<sup>18</sup>

Some generalists might resist this move. Generalists hold that perceptual contents are best characterized by existentially quantified statements (e.g., Byrne 2001, Hill 2009, 2021). The generalist might argue that general contents can indeed enable *de re* thoughts. Alternatively, they might simply deny that perception always enables demonstrative thoughts. Recently, Hill (2021) has defended a generalist account of *perceptual awareness*. Roughly, for Hill, one is perceptually aware of an object O when: O satisfies a set of features attributed to some object in perceptual experience, O causes the experience, and O is the best causal candidate for satisfying the general content (2021: 1392). It might be thought that perceptual awareness is enough to ground *de re* thoughts. On the other hand, one might just think that perceptual awareness is enough to explain

---

<sup>16</sup> The term ‘*de re*’ is used in different ways in the literature. I use it here to mean singular non-descriptive thinking.

<sup>17</sup> While this connection between perception and thought is widely assumed in contemporary contexts, it is not without its opponents. For example, Russell (1911, 1912) didn’t allow *de re* thoughts about perceptual objects due to his restrictive views on acquaintance (a necessary condition for such thoughts).

<sup>18</sup> It might be thought that viewing perceptual MOPs as *de re* requires a Russellian view of perceptual content. On this view, content constituents are composed of objects and properties themselves. However, while Fregean MOPs are often interpreted to involve descriptive contents, theorists have argued that there can also be *de re* MOPs, e.g., Evans (1982), Bach (1987), and Recanati (2012).

the practice of using demonstratives to describe perceptual experiences (see Hill 2021: 1404). I will not make any claims here to whether such generalist approaches are plausible. Rather, I've provided some prima facie motivation for thinking that a perceptual MOP must itself be *de re* in order to enable *de re* thoughts. However, I ultimately leave it as open-ended whether descriptive perceptual MOPs (i.e., MOPs that fix reference descriptively) could enable *de re* thoughts.

Let us take stock. In the last four sub-sections, I've argued that perceptual MOPs: meet Frege's criterion, are not intentional modes, are not individuated by the representation of distinct attributes, and enable *de re* thoughts.

## 2. Perceptual MOPs and object files

In this section, I defend the view that object files are token perceptual MOPs. Object files are psychological entities. They are the representations perceptual systems employ when they represent an object under some MOP. The MOPs that constitute Fregean perceptual contents can be understood as abstractions from these token representation instances. If this is the case, then an object file for some object *a* is metaphysically prior to the perceptual MOP for *a*. The abstraction from the token perceptual MOP for *a* depends on the object file instance.

The central aim of this paper is to provide an account of token perceptual MOPs. There are a variety of ways of understanding the relation between these token MOPs and some kind of abstracta. I have proposed one way to understand this relation. Nevertheless, I acknowledge that there are alternative ontological frameworks equally consistent with the principal claims made in this paper.<sup>19</sup>

On the proposed view, token perceptual MOPs are object files. This view has a precursor in the *mental files* (MFs) approach to Fregean thought contents (e.g., Bach 1987, Recanati 2012). MFs store a collection of properties that an agent believes the referent satisfies. The MF framework proposes an attractive account of Frege's puzzle at the level of thought. For example, in the case of Venus, the agent has initiated two files for the same referent. Once they realize that the MORNING STAR-FILE and the EVENING STAR-FILE co-refer, they can link the files together (Recanati 2012: 43, 197). In this framework, the reason that the agent is prone to have contradictory thoughts about the same object is that they have initiated two distinct co-referential MFs.

---

<sup>19</sup> For example, one might also consider perceptual MOPs as types or abstract entities. These abstract MOPs are part of a proposition. Object files would then be token perceptual MOPs of these types.

My proposal follows a similar line of thinking for perception. Object files are a theoretical construct currently used in cognitive science, and first developed by Kahneman, Treisman and Gibbs (1992). These files are representations that store information in visual short-term memory (VSTM) about objects in one's visual field, and this information is updated when sensory conditions change.<sup>20</sup> The referents of the object file representations are plausibly determined by an object's estimated spatiotemporal location (via what Pylyshyn calls 'FINSTs' (2007)), not by any conceptual or descriptive information about the object contained in the file (Kahneman et. al 1992: 178).<sup>21</sup> Cognitive agents are able to re-identify particular objects in vision by exploiting the object-specific information stored in these files.

There is significant empirical evidence that our perceptual system represents object files. For example, Kahneman et al. (1992) showed that the object files construct explains *object-specific preview benefits*. In tasks that exhibit this benefit, participants are shown a set of objects with *preview features* shown for a brief period of time (e.g., certain letters in the middle of square frames). Once the square objects change location on the screen, they are presented with either distinct features or the same preview features. When there is a match, the participants exhibit faster reaction times in identifying the objects. These preview benefits are explained by an object file representation where objects are tracked via spatiotemporal continuity and the preview feature is stored within the file (Kahneman et. al 1992, Noles et. al 2005, Kimchi and Pirkner 2014).

The existence of object files is also supported by *multiple object tracking* (MOT) experiments. In these tasks, participants need to keep track of a subset of target objects within a larger set of objects. All the presented objects move randomly with distinct trajectories around the screen. It has been shown that subjects can track roughly four objects in MOT experiments (Pylyshyn and Storm 1988, vanMarle and Scholl 2003). This suggests that perceivers can initiate roughly four object files in VSTM at any given time.

Other tasks that test VSTM provide evidence for object files. For example, Luck and Vogel (1997) had subjects keep track of colors and orientations of a set of objects. Participants were able to track a conjunction of features of objects just as well as a single feature of objects. The load on

---

<sup>20</sup> In a recent paper by Green and Quilty-Dunn (2020), the authors discuss what kind of representation object files are. According to their account, object file representations have a propositional, as opposed to iconic, format. Furthermore, they propose a 'multiple-slots' view according to which files store different perceptual features in distinct memory stores.

<sup>21</sup> For a recent discussion of this issue, see Quilty-Dunn and Green (2021). The authors defend a "flexible model" of attributional reference fixation. According to this view, object files *can* fix reference via attributional information. However, only for certain attributional information depending on contextual parameters. The authors' view of attributional reference fixation involves primarily the *maintenance* of reference through time, as opposed to the initial reference fixing (what they call 'locking').

VSTM was shown to depend, not on the number of features for objects held in VSTM, but rather the amount of *integrated object representations*. Thus, the study provides evidence for object files and indicates that there are roughly four available for any given perceptual episode.

On the proposed view, token perceptual MOPs are intermodal object files (see Jordan et. al 2010). When a perceiving agent has initiated two distinct object files for the same object, their perceptual system is in a position to falsely attribute inconsistent properties to the same object. In this case, perceptual experience might represent inconsistent properties as belonging to the same object. This happens when the perceptual system is unaware that distinct object files are co-referential. In this way, initiating distinct co-referential object files meet's Frege's criterion for a difference in MOP. In other words, the object files view explains why perceivers are in the special epistemic predicament outlined in section 1.1.

Thus, the object files account can explain the perceptual Frege cases noted above. For example, in the case of double vision, the perceiver initiates two distinct object files for the same object. They then place distinct spatial information into each file. Likewise for Perry's dog around the pole case. Two object files are initiated for the same dog. We can say that one file represents and stores the information " $o^1$  is to the right of the pole," and another file represents " $o^2$  is to the left of the pole."

For the crossmodal case specified in section 1.1, the proposed framework suggests that the agent has initiated two intermodal object files for the same object. In one file, the subject puts the property of being 'to the left of the pole' (acquired via vision) and in the other 'to the right of the pole' (acquired via audition). Other intermodal perceptual Frege cases can be explained in a similar fashion.

Hence, the object file framework has great explanatory power when applied to perceptual Frege cases. The object file view of perceptual MOPs clearly meets the first constraint. The initiation of distinct object files is a sufficient condition for perceptual Frege cases. If some agent S initiates two co-referential object files for  $a$ , then S is automatically prone to make inconsistent attribution representations regarding  $a$ . This will remain the case until the perceptual system becomes sensitive to the fact that the object files co-refer (see section 1.2). This suggests that the initiation of co-referential object files is a sufficient condition for perceptual Frege cases. Again, this is defined as an epistemic situation where the perceiver might attribute inconsistent properties to the same object.

The object files framework also clearly meets the second constraint. Object files are not intentional modes. Different intentional modes with the same content  $C$  do not represent distinct information to perceivers. However, the initiation of two distinct, but co-referential, object files does represent distinct information to perceivers. When a perceptual system initiates two co-referential object files, it is in a position to ‘learn something new’ when it becomes sensitive to the fact that the two object file representations are co-referential. In this vein, objects files corresponding with two MOPs  $a$  and  $b$ , e.g.,  $OF^a$  and  $OF^b$ , represents different information in perceptual experience because a perceptual system can be unaware, and then ‘learn,’ that  $OF^a$  co-refers with  $OF^b$ . Following the discussion in section 1.2, this means that, minimally, a perceptual system becomes sensitive to the fact that distinct object files are co-referential. This is the perceptual analog to being able to learn that  $a = b$  at the level of thought. Thus, distinct co-referential object files represent different information to perceivers and hence correspond with distinct content constituents. In this way, object files are not intentional modes.

My tentative suggestion in section 1.2 was that this ‘learning’ transition takes two forms. First, I suggested that ‘learning’ that two distinct MOPs are co-referential can result in two object representations becoming one. In this case, two object files become one file. Such a transition is the perceptual analog to what Recanati calls the ‘merging’ of MFs, where two MFs become one (2012: 44). Second, I suggested, following O’Callaghan (forthcoming), that the perceptual system can continue to represent two distinct objects (where two object files remain) but become sensitive to the fact that they are co-referential. This is roughly the perceptual analog to what Recanati calls the ‘linking’ of MFs, where two MFs remain but the subject recognizes that they are co-referential and can integrate information contained within them (see 2012: 43).

Object files also meet the third constraint. Object files are not themselves constituted by, or are individuated in virtue of, the representation of distinct properties attributed to an object. Rather, they specify how different features belong to various objects in one’s perceptual field. These files play an organizational or structural role in perception. This organizational role is to individuate objects and associate certain perceptual information with them and not to others. Object files are not individuated in virtue of the features contained within the files. Rather, they

are plausibly individuated via their initiation conditions. These are the conditions, usually some kind of perceptual relation to an object, such that the file is initiated.<sup>22</sup>

Lastly, the object file view shows a clear connection between the non-descriptive character of perception and thought. I assume that, following Pylyshyn (2007), the reference of object files is fixed via what he calls “visual indexes” or “finger instantiations” (FINSTs for short). A FINST is a mechanism of early vision that performs tracking of objects. FINSTs track and individuate objects purely in virtue of their spatiotemporal continuity.<sup>23</sup> They do not represent objects as falling under a particular conceptual category (2007: 56, 94). Following Pylyshyn, I will assume that FINSTs secure the referent of object files (2007: 38). On a plausible reading of Pylyshyn, FINSTs are distinct from object files. However, recent evidence suggests that FINSTs are likely best understood as a demonstrative *component* of object files (see Bahrami 2003, Haladjian and Pylyshyn 2008).<sup>24</sup> In other words, object files are plausibly complex representations that incorporate both the representation of features and a demonstrative-like representation of individuals (e.g., a FINST).

Either way, FINSTs involve a non-satisfactional mechanism of reference.<sup>25</sup> As Pylyshyn notes, FINSTs are “not conceptual, these sensory individuals are not represented *as* objects or as Xs for *any* possible category X. They are just picked out transparently by a causal or informational process without being conceptualized as something or other” (2007: 56). Pylyshyn’s treatment of FINSTs is not without its critics (e.g., see Burge 2010).<sup>26</sup> However, I am assuming that object files at least involve a demonstrative-like representation of individuals. This view is held by many, e.g., see Murez and Recanati (2016), Echeverri (2016) and Echeverri (2017). When object files are initiated, a subject’s perceptual system associates a cluster of information with *that*, i.e., some object in their perceptual field. This non-descriptive presentation of objects can be preserved in

---

<sup>22</sup> For present purposes, I’ve left this characterization of ‘initiation conditions’ abstract. For more detailed discussions, see Carey (2009) and Burge (2010), who both suggest that the object file system selects and tracks individuals that are three-dimensional, cohesive and bounded. By contrast, Green (2018) has recently argued that the object file system selects and tracks perceptual individuals when they satisfy a more permissive set of organizational principles. Pylyshyn has suggested that “a FINST provides a reference link from an object files to an object in the distal environment, and...it does so in response to a causal/informational event in the world that captures the FINST reference tokens” (2007: 39).

<sup>23</sup> Pylyshyn argues that a variety of empirical phenomena suggest that our visual system includes such a mechanism (see 2007: 22 – 29).

<sup>24</sup> For example, Bahrami (2003) used a combined MOT and change detection task. They showed that detection of feature changes was better for tracked objects. This suggested that FINSTs (which track objects) are components of object files. Thanks to Jake Quilty-Dunn for bringing these studies to my attention.

<sup>25</sup> I acknowledge that FINSTs are not the only possible perceptual mechanisms for non-satisfactional reference. I’ve only shown that, if we adopt Pylyshyn’s view of FINSTs, then object files meet the singular reference constraint.

<sup>26</sup> For example, Burge challenges the view that there can be FINST indexes that “are not accompanied by any representation that ‘encodes’ a property” (2010: p. 455). This criticism is an instance of Burge’s view that ‘bare’ perceptual reference does not occur. For Burge (2010), perceptual reference must involve some minimal form of perceptual attribution. In this vein, Burge might reject the possibility of *de re* MOPs as I’ve characterized them here (since they are *de re* in virtue of utilizing non-satisfactional mechanisms of reference).



thought such that *de re* thoughts, including demonstrative thoughts, are enabled. In other sensory modalities outside of vision, a demonstrative component of object files likely fixes the reference for intermodal files. In this way, the object files view also meets the fourth constraint.<sup>27</sup>

Hence, on the proposed view, perceptual MOPs are abstractions from object file instances. A token perceptual MOP is an object file representation in some perceptual system. For any object there is a limited range of possible perceptual MOPs for that object. This limited range of MOPs corresponds with the limited amount of object files that can be represented in VSTM during some perceptual episode. In most experimental contexts, this has been shown to be roughly four (Pylyshyn and Storm 1988, Luck and Vogel 1997, vanMarle and Scholl 2003). The object component of Fregean perceptual content is thus determined by two factors: 1) what object the system is perceptually related to and 2) which object file representation(s) have been initiated for that object.

To sum up, I've argued that the object file view of perceptual MOPs satisfies all four constraints. Object files meet Frege's criterion, they are not intentional modes, they are not individuated in virtue of the representation of distinct attributed features, and they enable *de re* thinking. Object files are thus plausibly token perceptual MOPs.

### **3. Alternative accounts of perceptual MOPs**

In *The Unity of Perception* (2018), Schellenberg understands a variety of perceptual phenomena through a framework called *capacitism*. Within this framework, perceptual states are constituted by the employment of various perceptual capacities. These capacities function to discriminate and single out mind-independent particulars, i.e., objects, property-instances and events (2018: 29). Schellenberg suggests that perceptual capacities are the perceptual counterparts to Fregean MOPs. While Fregean MOPs can be understood as various ways of thinking about, or referring to, the same referent, each perceptual capacity is a distinct way in which we single out the same type of particular (2018: 52, 85). Perceptual MOPs for Schellenberg can differ in two ways. First, they can differ with respect to the perceptual capacities that are employed. Second, they can differ with respect to which particular the capacities single out (2018: 85-89).

---

<sup>27</sup> As an additional motivation for this point, several theorists have argued that thinking of an individual through a MF is constitutive of *de re* thought (Jeshion 2001, Jeshion 2010, Recanati 2012). Hence, the view that something like an object file enables *de re* thinking has precedence in the MF literature.

For Schellenberg, different perceptual capacities that single out the same object correspond with different MOPs for that object. Consider an example she provides where a subject singles out a triangle in different ways, either via its 3 sidedness or its 3 corneredness (2018: 52). These distinct ways of singling out triangles correspond with the employment of distinct perceptual capacities, which would in turn correspond with distinct, but co-referential, perceptual MOPs. Employing these two distinct perceptual capacities produce slight variations in phenomenal character.

A problem with Schellenberg's view is that it cannot adequately explain some perceptual Frege cases. Consider again the double vision case. Remember, in this case some object *a* appears to you twice (as *a* and *b*). You represent a spatial relation between *a* and *b*. In this case, there is no reason why the perceiver cannot employ the *same perceptual capacity* for singling out both *a* and *b*. Schellenberg's framework does not rule out the possibility that we can employ the same capacity for distinct instances of particulars presented simultaneously. For example, if *a* and *b* are the same triangle, the capacity for discriminating triangles via its 3-sidedness might be employed twice (for both *a* and *b*). But if this is the case, then Schellenberg's framework cannot explain why such a perceptual state would constitute a Frege case. Perceptual Frege cases require the representation of distinct but co-referential MOPs. However, on Schellenberg's framework, it is possible that this case involves the presentation of the same MOP represented twice.

If the same MOP is represented twice, then a rational perceiver cannot attribute inconsistent properties to it. So, according to Schellenberg's framework, we would have to conclude that, if the same perceptual capacity is employed twice in the case of double vision, then the perceiver in this case is not rational. But this is counterintuitive. Given the circumstances of intoxication and double vision, the agent should not be deemed irrational for perceptually attributing inconsistent properties to the same object. These considerations suggest that this case cannot be adequately explained by Schellenberg's framework. A similar worry can be raised for the other Frege cases mentioned in section 1.1 that involve attributing inconsistent spatial properties to objects.

Schellenberg's account plausibly captures intentional modes better than perceptual MOPs. When a subject singles out the same object via distinct perceptual capacities, these constitute two different ways in which the subject relates to the same content. Like the employment of distinct capacities for the same object, different intentional modes with the same content also consequence in distinct phenomenological experiences. However, if we understand Schellenberg in this way,

then the employment of perceptual capacities would constitute fine-grained perceptual modes. The employment of the capacity for singling out triangles via their 3-sidedness, for example, would present content involving triangles in a very specific perceptual manner of representation. This specificity cannot be captured by, for example, merely a ‘visual’ or ‘haptic’ intentional mode.

Another theory of perceptual MOPs is defended by Chalmers (2004). Chalmers defends a theory of descriptive (sometimes called ‘*de dicto*’), as opposed to *de re*, perceptual MOPs. The plausibility of descriptive perceptual MOPs likely depends on views regarding whether perceptual content should be construed as general or particular. I do not endorse either view in the present article.<sup>28</sup> Putting the generalist and singularist debate aside, I show that Chalmers’ view is at least inconsistent with the third constraint defended in section 1.3.

Chalmers proposes that perceptual MOPs are *conditions on extension*. These conditions specify what must be true of the object in order for it to be the referent of the perceptual MOP (2004: 171). For example, when you perceive a green sphere, the Fregean content associated with that sphere might be the condition “*the object that is causing this experience in the appropriate way*” (2004: 173). Perceptual Frege cases involve the presentation of multiple co-referring objects. Thus, a perceptual MOP, understood as a condition on extension, will involve a way to individuate distinct co-referential MOPs that are simultaneously presented in experience. For example, going back to our double vision case, a perceptual MOP for *a* might be characterized by *the object causing my experience of the object that is to the right of the streetlight*. A MOP for *b* might be characterized by *the object causing my experience of the object that is to the left of the streetlight*.<sup>29</sup> In this way, for Chalmers, two co-referential MOPs are individuated in virtue of distinct satisfaction conditions which involve the attribution of a distinct set of features to the same object.

As noted in section 1.3, perceptual MOPs plausibly should not be individuated via the representation of distinct features that are attributed to the same object. This follows a popular view about MOP individuation that I recognize some theorists will reject. Nevertheless, if this individuation constraint is accepted, Chalmers’ view fails to meet it. Different MOPs, understood as distinct conditions on extension, are individuated in terms of distinct satisfaction conditions. In

---

<sup>28</sup> For some recent discussions, see Schellenberg (2018) and Hill (2021).

<sup>29</sup> For Chalmers, perceptual MOPs correspond with what he calls *primary intensions* of linguistic expressions (Chalmers 2004, 2006). Primary intensions are functions from possible worlds to extensions. If the primary intension of ‘water’ can be characterized roughly as ‘the drinkable liquid found in oceans and lakes,’ then the primary intension of ‘water’ maps our world, a world in which H<sup>2</sup>O meets this condition, to H<sup>2</sup>O. Alternatively, it would map the Twin Earth possible world to XYZ. This is the case since in this other possible world, it is XYZ, not H<sup>2</sup>O, that meets the condition specified by the primary intension (Chalmers 1996: 57, 2006: §2.1).

the case of perceptual objects, this will involve distinct conditions on extension analyzed by the attribution of distinct features to some object. While plausibly less salient than the first two constraints specified in section 1.1 and section 1.2, I've motivated reasons for accepting this constraint that Chalmers' view fails to meet.

In section 1.4, I primarily argued that perceptual MOPs must enable *de re* thoughts, not that they themselves must be *de re* MOPs. As already mentioned, there might be generalists who deny that enabling *de re* thoughts implies that the perceptual MOP itself must be *de re*. Depending on whether you think descriptive MOPs can enable *de re* thoughts, Chalmers' view can be seen to either meet or fail to meet this fourth constraint. If it is thought to meet the constraint, it must be explained how perceptual MOPs can have satisfactoral mechanisms of reference while their corresponding cognitive MOPs have non-satisfactoral ones. Chalmers' conditions on extension, in the case of perception, demonstratively refer to experiences. However, in perceptual Frege cases, they use satisfactoral mechanisms to fix reference for the (apparently) distinct objects presented in that experience.<sup>30</sup>

## Conclusion

In this paper, I've outlined four constraints that I've argued theories of perceptual MOPs should adhere to. I defended the view that token perceptual MOPs are object files. On the proposed view, when a subject's perceptual content has two distinct MOPs for *a* as constituents, then the subject initiates two distinct object files that both refer to *a*. Perceptual MOPs are abstractions from token perceptual MOPs. These token perceptual MOPs are object file representations in perceptual systems. My defense of the proposed framework appealed to its satisfaction of the four constraints outlined in section 1. Furthermore, I've argued that Schellenberg (2018) and Chalmers' (2004) accounts fail to meet some of these constraints.<sup>31</sup>

---

<sup>30</sup> This is not meant as an exhaustive treatment of all existent theories of perceptual MOPs. For example, I have not evaluated or compared Burge's treatment of perceptual MOPs as distinct 'perceptual perspectives' on particulars (2010: 385, 411). Rather, I've limited my analysis to two discussions that have provided special attention to perceptual MOPs of objects. In addition, as one reviewer points out, Burge's (2022) just-released book raises intriguing questions regarding how some of Burge's views relate to those developed in the present article. Due to space limitations, I regret that I am unable to include a full comparison with the views described in Burge's new book.

<sup>31</sup> I am grateful to Casey O'Callaghan for comments on multiple drafts of this paper. Comments from Jake Quilty-Dunn and three anonymous reviewers also led to significant improvements. I also thank Judith Carlisle, Ron Mallon and Emily Prychitko for helpful comments on earlier drafts.

## References

- Bach, Kent (1987). *Thought and reference*. Clarendon Press.
- Bahrami, B. (2003). Object property encoding and change blindness in multiple object tracking. *Visual Cognition*, 10(8), 949–963.
- Bourget, David (2017). Representationalism and Sensory Modalities: An Argument for Intermodal Representationalism. *American Philosophical Quarterly* 54 (3): 251-268.
- Burge, Tyler. (2010). *Origins of objectivity*. Oxford University Press.
- \_\_\_\_\_ (2022). *Perception: first form of mind*. Oxford University Press.
- Byrne, A. (2001). Intentionalism defended. *The Philosophical Review*, 110, 199–240.
- Carey, Susan. (2009). *The origin of concepts*. Oxford University Press.
- Chalmers, David J. (1996). *The conscious mind: In search of a fundamental theory*. Oxford Paperbacks.
- \_\_\_\_\_ (2004). The representational character of experience. In Brian Leiter (ed.), *The future for philosophy*, pp. 153-181. Oxford University Press.
- \_\_\_\_\_ (2006). Two-dimensional semantics. In Lepore, Ernest & Smith, Barry C. (eds.), *Oxford handbook of the philosophy of language*, pp. 574-606. Oxford University Press.
- Crane, T. (2003). The Intentional Structure of Consciousness. In Quentin Smith and Aleksandar Jokic (Eds.), *Consciousness: new philosophical perspectives*, pp. 33–56. Oxford University Press.
- Echeverri, S. (2016) "Object Files, Properties, and Perceptual Content", *Review of Philosophy and Psychology*, 7/2, 2016, 283-307.
- \_\_\_\_\_ (2017) "Visual Reference and Iconic Content", *Philosophy of Science*, 84/4, 2017, 761-781.
- Evans, G. (1982). *The varieties of reference*. Oxford University Press.
- Frege, Gottlob. (1892). On Sense and Nominatum. Ed. Aloysius P. Martinich. *The philosophy of language*, pp. 221-27. Oxford University Press, 2001.
- \_\_\_\_\_ (1918). Thought. *The Frege reader*, by Michael Beaney, pp. 325–345. Wiley, 1997.
- Green, E. J. (2018). What do object files pick out? *Philosophy of Science*, 85(2), 177-200.
- \_\_\_\_\_ & Quilty-Dunn, J. (2020). What is an object file? *The British Journal for the Philosophy of Science*.

- Haladjian, H. H., & Pylyshyn, Z. W. (2008). Object-specific preview benefit enhanced during explicit multiple object tracking. *Journal of Vision*, 8(6), 497.
- Hill, C. S. (2009). *Consciousness*. Cambridge: Cambridge University Press.
- \_\_\_\_\_ (2021). Perceptual existentialism sustained. *Erkenntnis*, 86(6), 1391-1410.
- Jeshion, Robin. (2001). Acquaintanceless De Re Belief. *Truth and meaning: investigations in philosophical semantics*. New York: Seven Bridges Press.
- \_\_\_\_\_ (2010). Singular Thought: Acquaintance, Semantic Instrumentalism, and Cognitivism. *New essays on singular thought*. Oxford University Press.
- Jordan, K. E., Clark, K., and Mitroff, S. R. (2010). “See an Object, Hear an Object File: Object Correspondence Transcends Sensory Modality.” *Visual Cognition* 18: 492–503.
- Kahneman, D., Treisman, A., & Gibbs, B.J. (1992). The reviewing of object files: Object specific integration of information. *Cognitive Psychology* 24, 175–219.
- Kaplan, D. (1989). “Demonstratives,” in J. Almog, J. Perry and H. Wettstein (eds), *Themes from Kaplan*, pp. 481–564. Oxford University Press,
- Kimchi, R. and Pirkner, Y. (2014). ‘Surface Feature Congruency Effects in the Object- Reviewing Paradigm Are Dependent on Task Memory Demands’, *Psychonomic Bulletin and Review*, 21, pp. 1019–25.
- Quilty-Dunn, J., & Green, E. J. (2021). Perceptual attribution and perceptual reference. *Philosophy and Phenomenological Research*.
- Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, 390(6657), 279-281.
- Millar, B. (2016). Frege’s Puzzle for Perception. *Philosophy and Phenomenological Research* 93(2), 368–392.
- Murez, Michael, and François Recanati. (2016). Mental files: An introduction. *Review of Philosophy and Psychology* 7:265–281.
- Noles, N. S., Scholl, B. J. and Mitroff, S. R. (2005). The Persistence of Object File Representations. *Perception and Psychophysics*, 67, pp. 324–34.
- O’Callaghan, Casey. (2014). Intermodal binding awareness. In D. Bennett, & C. Hill (eds.), *Sensory integration and the unity of consciousness*, pp. 73-104. MIT Press.
- \_\_\_\_\_ (2019). *A multisensory philosophy of perception*. Oxford University Press.

- \_\_\_\_\_ (forthcoming). Crossmodal Identification. In Mroczko-Wąsowicz, A. & Grush, R. (Eds.). *Sensory individuals: contemporary perspectives on modality-specific and multimodal perceptual objects*. Oxford University Press.
- Perry, J. (2001). *Reference and reflexivity*. Stanford: CSLI.
- Pylyshyn, Zenon W. (2007). *Things and places: how the mind connects with the world*. MIT Press.
- Pylyshyn, Z. W., & Storm, R. W. (1988). Tracking multiple independent targets: Evidence for a parallel tracking mechanism. *Spatial vision*, 3(3), 179-197.
- Recanati, François (2012). *Mental files*. Oxford University Press.
- Rescorla, M. (2020). Perceptual co-reference. *Review of Philosophy and Psychology*, 11(3), 569-589.
- Richard, A. M., Lee, H., & Vecera, S. P. (2008). Attentional spreading in object-based attention. *Journal of Experimental Psychology: Human Perception and Performance*, 34(4), 842.
- Russell, Bertrand. (1911). Knowledge by Acquaintance and Knowledge by Description. *Proceedings of the Aristotelian Society* 11.1: 108-28
- \_\_\_\_\_ (1912). *The problems of philosophy*. New York: Oxford University Press, 1997.
- Thompson, B. (2009). Senses for senses. *Australasian Journal of Philosophy*, 87(1), 99-117.
- Treisman, A. (1998). "Feature binding, attention and object perception." *Philos Trans R Soc Lond B Biol Sci*. Aug 29; 353(1373): 1295–1306.
- Schellenberg, Susanna (2018). *The unity of perception*. Oxford University Press.
- Speaks, Jeff. (2015). *The phenomenal and the representational*. Oxford University Press.
- vanMarle, K. and Scholl, B. J. (2003). Attentive Tracking of Objects Versus Substances. *Psychological Science*, 14, pp. 498–504.