PERCEPTUAL ATTRIBUTION AND PERCEPTUAL REFERENCE

Jake Quilty-Dunn
Washington University in St. Louis

E.J. Green
Massachusetts Institute of Technology

Forthcoming in *Philosophy and Phenomenological Research*

Abstract: Perceptual representations pick out individuals and attribute properties to them. This paper considers the role of perceptual attribution in determining or *guiding* perceptual reference to objects. We consider three extant models of the relation between perceptual attribution and perceptual reference—all attribution guides reference, no attribution guides reference, or a privileged subset of attributions guides reference—and argue that empirical evidence undermines all three. We then defend a *flexible-attributives model*, on which the range of perceptual attributives used to guide reference shifts adaptively with context. This model underscores the remarkable and dynamic intelligence of our perceptual capacities. We elucidate implications of the model for the boundary between perception and propositional thought.

1. Introduction

Imagine that a chameleon enters your field of vision. Its unusual shape and bright chartreuse skin attract your attention. Then, frightened by your presence, it shifts to a dull seaweed green. While your visual experience updates to register the color change, the fact that a single enduring individual sported both shades of green is reflected in your experience as well. Thus perception distinguishes the individuals perceived from the properties they apparently manifest.

More technically, it seems that perceptual representation comprises at least two aspects: *reference* to perceptible individuals, and *attribution* of properties to those individuals. The distinction between perceptual reference and perceptual attribution is widely accepted in philosophy (Smith 2002; Clark 2004; Cohen 2004; Burge 2010a; Green 2017; Rescorla 2018; Millar forthcoming; Lande unpublished; cf. Fodor 2007; Schellenberg 2018, 67-68; 2020) and psychology (Treisman & Gelade 1980; Kahneman et al. 1992; Scholl & Leslie 1999; Pylyshyn 2007; cf. Spelke 1988). It is presently unclear how to distinguish these aspects of perceptual representation. For example, are there distinct constituents of the vehicles of perceptual representation responsible for reference and attribution as in a sentence like 'This is green', or are percepts structured more like images, with individuals and their properties represented by means of the same parts of the representation? It is also unclear how reference and attribution interact in perception. In particular, one may wonder how attribution constrains or informs perceptual reference. Is perceptual reference to an individual determined independently of the properties attributed to it, or does perceptual reference depend on the contents of perceptual attribution? This question will be the focus of this paper.

A full understanding of perceptual representation requires understanding the relation between perceptual attribution and perceptual reference. This relation arguably has consequences for the sort of attribution achievable in perception. Burge (2009; 2010a; 2010b) holds that perceptual representation lacks predicative propositional content as exhibited by sentences like 'This is F', instead possessing content akin to a complex demonstrative like 'That F'; i.e., perception involves the *de re* application of attributives rather than predication. For Burge, this distinction is grounded in the fact that perceptual attribution occurs within the scope of a referential noun-phrase-like structure and thus functions to *guide* singular perceptual reference, while genuine predication does not have this constitutive guiding function. Thus, answering the question whether perceptual attribution always serves to guide perceptual reference is critical for determining whether Burge's non-propositional model of perceptual attribution is sustainable. A negative answer to this question raises doubts about Burge's *That F* model and opens up the possibility that some instances of perceptual attribution constitute genuine predication (Quilty-Dunn forthcoming).

We assume that perception involves singular reference rather than, for example, existentially quantified contents. As will become clear below, object perception is best explained by appeal to referential constituents of perceptual representations. We leave it open whether other aspects of perception may have quantification-involving contents (Hill 2019; cf. Burge 2010b). Our focus will be the relation between perceptual reference to external objects/individuals and perceptual attribution of properties to those individuals.[1]

Object perception provides a useful case study in the interaction between perceptual attribution and perceptual reference. Perceptual object representations ("PORs") constitute a natural representational kind (Smortchkova & Murez forthcoming) and have received intensive empirical study. PORs track individuals over time while simultaneously maintaining and updating property-attributing elements ("attributives"). Philosophers and cognitive scientists have constructed sophisticated theories of PORs concerning how PORs refer, how they encode properties, and how their reference does or doesn't depend on the properties they encode. As we show in Section 2, theorists have defended positions that variously suggest that the referential aspect of PORs depends on *no* attributives, *all* attributives, or a *privileged subset* of attributives.

[1] We set aside *de se* reference and egocentric indexing (cp. Burge 2009, 256).

Taking object perception as our case study, we will argue in Section 3 for a *flexible model* of the relation between attribution and reference. Specifically, the way the visual system maintains perceptual reference to objects depends on *some* but *not all* attributive elements of perceptual content, and which attributives are used in determining reference shifts depending on context in a systematic, adaptive fashion. Thus, not all elements of perceptual content should be modeled on Burge's *That F* structure, nor is perceptual reference secured by means of a privileged class of attributives or in a manner that is unmediated by attribution altogether. This flexible model of perceptual reference emphasizes the remarkable and dynamic intelligence of our perceptual capacities.

2. Three Models of the Perceptual Reference-Attribution Relation

Our central case study in the interplay between perceptual attribution and perceptual reference will be perceptual object representations ("PORs"), also known as *object files*. Before surveying models of the reference-attribution relation, we'll describe the basic functional properties of PORs and the experimental paradigms used to study them.

2.1—Object perception. PORs represent a small number of perceptible individuals, encode and store information about them, and are used to perceptually track those individuals over time and space. Various experimental paradigms probe the referential and attributive contents of PORs through certain functional signatures. These paradigms include:

Multiple object-tracking: Participants are shown a display containing some number of items (e.g., 8), and the experimenters indicate that some subset (e.g., 4 flashing items) are to be tracked (Pylyshyn & Storm 1988). Participants are typically effective at tracking up to about 4 items (Pylyshyn 2003), and substantially more if the objects move slowly (Alvarez & Franconeri 2007). Subjects are superior at detecting changes in shape and color of tracked vs. untracked items (Bahrami 2003), suggesting that representations used for tracking also attribute properties to tracked individuals.

Preview benefit: Subjects see features appear on visible objects (e.g., a letter appearing within each item's boundary), the features disappear, and the objects move to new locations. Then a feature appears in one object and subjects perform an object-reviewing task, e.g., indicating whether it matches one of the original features (Kahneman et al. 1992; Noles et al. 2005). Performance is quicker and more accurate when previewed features re-appear in the same object, despite the objects having moved to new locations. This *object-specific preview benefit* exploits a representation that picks out individual objects, tracks them over time and to new locations, and stores information about their properties, including past properties. When a new feature matches information already stored in the POR for the object in which it appears, responses to the feature are speeded.

Transsaccadic memory: Subjects are instructed to move their eyes to or near an object. During the eye-movement or "saccade", the object may move, change its features, or remain unchanged, and subjects are instructed to move their eyes to its new location (Hollingworth et al. 2008; Richard et al. 2008), discriminate some feature (Pollatsek et al. 1984; Irwin 1992), or perform a related task. The timing and accuracy of corrective saccades, feature naming/discrimination, and other tasks is modulated by a POR that is opened prior to (and survives during) the saccade. That is, PORs are constituents of "transsaccadic memory." They are used by the visual system to maintain a representation of enduring objects across changes in retinal stimulation caused by eye-movements.

These are three representative signatures of PORs. They share features, such as a characteristic typical limit of roughly 4 items. They also interact: the preview benefit is enhanced during multiple-object tracking (Haladjian & Pylyshyn 2008), previously fixated objects facilitate corrective saccades despite changes in location and color (Schut et al. 2017), and many transsaccadic-memory experiments are simply transsaccadic versions of the object-reviewing paradigm (Gordon et al. 2008). Other experimental paradigms that arguably probe PORs include bouncing/streaming displays (Meyerhoff & Scholl 2018), causal-launching perception (Moore et al. 2020), apparent motion (Odic et al. 2012), Ternus displays (Hein & Cavanagh 2012), the tunnel effect (Flombaum & Scholl 2006), and even enumeration of small sets (Chesney & Haladjian 2011), as well as a multiplicity of paradigms in developmental psychology (Carey 2009). PORs also seem to be the vehicles of object-based representation in visual working memory (Hollingworth & Rasmussen 2010; Gao et al. 2011; Green & Quilty-Dunn 2017; Quilty-Dunn 2019). There is also evidence that PORs are formed in non-visual modalities and may combine attributives from vision, audition, and touch (Zmigrod et al. 2009; Green 2019).

The presence of a natural representational kind is suggested by such "characteristic psychological effects" (Smortchkova & Murez forthcoming; cp. Carey 2009), and the fact that they cluster together provides evidence for a single underlying representation. This representation contains an element that referentially picks out an individual and attributives that encode features along dimensions like color, shape, orientation, and size, as well as categorical information (Green & Quilty-Dunn 2017). The fact that PORs succeed in representing individuals despite featural changes (Zhou et al. 2010) or even loss of featural attributives (Scholl et al. 1999; Bahrami 2003) implicates a robust reference/attribution distinction. The referential element of a POR can be maintained while attributives contained in the POR are discarded or modified.

Considerable research on PORs examines their role in solving the "correspondence problem" (Ullman 1979; Kahneman et al. 1992). Suppose you see an object (O1) at some location (L1). If the object moves to a new location (L2), the visual system must determine whether to update the location for O1 to L2, or instead to construct a POR for a novel object (O2). This process of determining whether a currently seen object *corresponds* to a previously seen object is a process of determining (and, when correspondence succeeds, sustaining) perceptual reference.

Object correspondence is an instance of the *maintenance* of perceptual reference over time. Reference maintenance is distinct from what we'll call reference *locking*, which is the process whereby reference to an object is initially secured. The principles governing these reference-determining mechanisms may diverge. To use a toy example, suppose initial reference locking requires that an object be visible and clearly distinguished from its surround (cf. Gao & Scholl 2010). It doesn't follow that reference maintenance shares these requirements—reference can be maintained even when objects disappear behind occluders (Scholl & Pylyshyn 1999; Flombaum et al. 2008). Moreover, it is possible that perceptual attribution sets different sorts of constraints on reference locking and reference maintenance. We can distinguish the features used to pick out an object in the first place from the features used to determine how the object persists over time. Green (2017) argues, for instance, that attribution of structural shape properties sets constraints on reference locking, since it is needed to distinguish perceptual reference to parts from perceptual reference to whole bodies. However, this view does not entail that shape information is always used in reference maintenance. Thus, reference maintenance is distinct from reference locking, though both are elements of the overall function of reference determination.

Reference maintenance sets constraints on viable views of the relation between perceptual reference and perceptual attribution. If, for example, continuity in certain perceptual attributives is irrelevant to reference maintenance, that puts pressure on the view that those attributives guide perceptual reference (pressure that could in principle be relieved by showing, e.g., that they guide reference *locking*). We will therefore look to the experimental literature on object correspondence to determine the role attributives play in reference maintenance, allowing us to draw abductive inferences about the role of attribution in perceptual reference more broadly.

2.2—Three models. When we consider whether perceptual reference depends on attribution, two extreme views present themselves: reference depends on *all* attributives, and reference depends on *no* attributives. Call these the *all-attributives* model and the *no-attributives* model, respectively. Although matters of interpretation are not straightforward, there is some textual basis for tentatively ascribing the all-attributives model to Burge and the no-attributives model to Pylyshyn. A third, intermediate position holds that perceptual reference depends on a special subset of attributives—call this the *privileged-attributives* model. We'll survey these views before arguing against them in favor of a *flexible-attributives* model, on which perceptual reference depends on some but not all attributives, and which attributives guide reference shifts depending on context.

We'll now clarify the notion of "guiding" reference, or the sense in which reference can or cannot "depend" on attribution (we use these terms interchangeably). We have in mind the phenomenon wherein *a perceptual process that has the function of locking or maintaining the reference of a perceptual representation R to some object relies on properties that R attributes (or recently attributed) to the object*. As Burge puts it, an attributive guides reference when it "is used by the individual or his representational system as an important restriction on the singular representation's purported referent" (2009, 257).

Consider reference locking. If R purports to single out an object and attributes some shape property S to it, then to say that the attributive guides reference is to say that the perceptual processes responsible for securing reference take into account that the object putatively has S. In veridical cases, the object actually has S; but the attributive can guide reference even when attribution fails to be accurate, in virtue of the attributive's figuring in the reference-determination process (cp. Campbell 2002, 93-94). A no-attributives model of reference locking would deny that any attributives in R figure in the processes that ground R's reference. Similarly, in deciding whether the object R referred to a moment ago is now present at a different location, reference-maintenance processes could take as input *all* the attributive elements of R (all-attributives model), *none* of them (no-attributives model), or *some* of them (privileged-attributives model; flexible-attributives model).

Our aim here is not to provide a definition or theory of reference guidance. However, we will make a key assumption about the function of guidance: If an attributive guides perceptual reference to some object, then the processes responsible for determining perceptual reference to that object should consult that attributive in their computational operations. This functional principle allows us to derive empirically testable hypotheses from claims about the guidance of perceptual reference. Specifically, claiming that some attributive guides the reference of some perceptual representation R commits one to the empirical claim that the attributive is used in processes that determine (lock or maintain) R's reference.

2.2.1—All-attributives. We begin with the all-attributives model. On this view, all perceptual attributions of properties to some external particular guide reference to that particular. In the case of reference locking, this view requires that, in veridical and referentially successful perception, the reference of a perceptual representation R to some object depends on every attributive element of R, and the object has every property R attributes to it. This description applies only to perfectly successful cases, however. We stress again that perceptual reference to an object can succeed even if a reference-guiding attribution is non-veridical. What matters is that the operations the visual system undergoes to lock the reference of R to the object consult all attributives. Similarly, in reference maintenance, the all-attributives model entails that processes that maintain R's reference to the object consult all attributive elements of R.

Burge holds that "[e]very perceptual attributive accompanies and guides a contextual singular perceptual application" (2010a, 539). For Burge, the crucial fact about perceptual attribution is not simply that it ascribes a property, but that its doing so subserves the more fundamental function of determining perceptual reference:

> In perception, this attributive function always serves and is subordinate to the larger perceptual function of identificational reference. The sole representational function of perceptual attributives is to effect an attribution as part of an identificational reference. Their scope is governed by a demonstrative-like, context-bound referential application: they are applied *by* the referring element.
>
> (Burge 2010b, 41)

We interpret these passages as endorsing the claim that every perceptual attributive functions to guide perceptual reference. For Burge, the fact that "all perceptual attributives are...applied within the scope of a context-bound, identificational, referential application" (*ibid.*) distinguishes them from predicative concepts and grounds the distinction between genuine predication and the sort of non-predicative attribution achievable in perception. Burge contends that, because perceptual attribution invariably functions to guide reference, perceptual content has a demonstrative-like *That F* structure. Propositional thought, on the other hand, can exhibit a predicative *That F is G* structure. This structure incorporates an attributive *G* that operates predicatively outside the scope of the referential noun phrase and (for Burge) therefore plays no reference-guiding role. Thus the role perceptual attribution plays in guiding reference lies at the core of the perception/thought distinction for Burge. Perceptual representations are constitutively non-propositional—and thus distinct from the predicative forms that characterize propositional thought—because they are limited to reference-guiding attributions rather than "pure predication" (Burge 2010b, 45).

Burge (2009, 2010b) emphasizes that perceptual attributives may guide reference without being veridical of the object to which reference is secured. Notice that any proponent of the all-attributives view *must* grant this possibility to avoid obvious counterexamples (e.g., perceptually referring to an object despite misrepresenting its color). Burge does suggest, however, that there are limits to how mistaken a perceptual representation can be while still securing reference. Specifically, he holds that for perceptual reference to "bodies" to succeed, the visual system must accurately attribute some generic shape information (e.g., topological boundedness or cohesion) to the referent (Burge 2009: 291-292).

As mentioned above, we assume that reference guidance has a functional signature: If an attributive guides reference to a particular, then it should be used by computations that subserve the locking and maintenance of reference to that particular. Thus the constitutive constraints Burge places on perceptual attribution appear to entail that all perceptual attributives should be consulted in reference locking and maintenance—i.e., the core claim of the all-attributives model.

We pause to note that Burge himself may reject this inference. Since he takes guidance to be primitive (2009, 257), he may deny that an attributive must play a certain functional role in perceptual processing in order to guide reference; or he may argue that an attributive may guide reference by contributing to some other reference-determining mechanism besides locking and maintenance. However, he also writes that "[g]uidance has to do with the functional importance that the representational system must *accord* a general representation in determining the purported referent of the context-bound singular representation" (Burge 2009, 258; emphasis his), suggesting that functioning in operations like those that subserve reference maintenance and locking is constitutive of guidance. In any case, the all-attributives model is an independently interesting proposal that follows from one natural interpretation of Burge, and is thus worth considering even if Burge would ultimately reject it.[2] It is also important to distinguish the claim that all instances of perceptual reference are guided by some attribution from the claim that all perceptual attributions serve to guide reference. While we will reject the latter claim, the account we'll defend is consistent with the former.

As we'll understand it, the all-attributives model claims that all perceptual attributives serve to guide perceptual reference. Given the functional signature of reference guidance, this implies that all attributives within a POR are in fact used by operations responsible for determining reference for that POR. To preview, we argue that this is not true because there are visual contexts that preclude the use of certain perceptual attributives in reference determination. A weaker view would hold merely that all perceptual attributives *can* serve to guide reference, or that some token attributives guide reference, even if other tokens of the same type do not. While not all perceptual attributives are actively *used* in reference determination, all are at least *usable*, insofar as they (or another token of the same attributive type) could be used for reference determination in some context. The problem with this model is that it cannot play Burge's advertised role in demarcating perceptual content from the content of propositional thought. As Burge himself observes (2010b, 41), the attributives employed in propositional thought are also usable in reference determination, and are so used when placed within a complex demonstrative. If I can think THAT TABLE IS BROWN, I could just as well use another token of BROWN to think THAT BROWN THING IS A TABLE, and thus use it to guide the reference of a demonstrative. So the weak view that all perceptual attributives are

[2] Another interpretive wrinkle arises in Burge's discussion of perceptual reference to "bodies" (roughly, bounded, cohesive 3D objects). As noted earlier, it seems to follow from Burge's view that when a property is perceptually attributed to an object, the attributive that represents the property serves to guide (lock or maintain) reference to the object. However, in certain places Burge seems to suggest the opposite. In discussing visual tracking capacities, Burge (2010a, 462-463) claims that "the visual system of human infants and many non-human animals are relatively impervious to changes in shape, color, ordinary kind (duck to sphere) in tracking bodies (whether moving or not), if continuity of cohesion and boundedness is maintained." This suggests that at least in certain contexts, attributives for determinate shape and color do *not* serve to guide tracking, and hence do not contribute to reference maintenance (see also Burge 2009, 257). It is unclear how to reconcile this claim with the view that "[e]very perceptual attributive accompanies and guides a contextual singular perceptual application" (Burge 2010a, 539). One option would be to claim that attributives for determinate shape and color only contribute to reference locking, not reference maintenance. Another option would be to claim that only a privileged class of spatiotemporal properties guides reference to bodies/objects (cp. Burge 2009, 280-281). We'll consider these options below. For now, we simply flag that we are unsure exactly which view Burge holds.

merely usable in reference determination (or that some tokens in some contexts are used in reference determination) is far less philosophically interesting than the view that all perceptual attributives really *do* play this role, barring malfunctions. Thus, we'll focus on the stronger view here.

Finally, we interpret the all-attributives view as a claim about the way in which our perceptual systems determine reference to feature-bearers. When we perceptually attribute redness to an object O, the attributive red serves to guide reference to O. One might, however, concede that some attributives fail to guide reference to objects, but insist that when this occurs the attributive instead guides reference to some distinct particular, such as a trope or property instance. So, the attributive red might not guide reference to any object, but it still guides reference to a redness trope.[3] One attributive would be responsible for two feature attributions: one to an object, one to a trope. Our focus is not on tropes, but we make two quick points here. First, even if attributives do guide reference to tropes, they also attribute properties to feature-bearers. We can therefore still ask whether attribution of features to feature-bearers also invariably guides reference to those feature-bearers. If not, the view that all perceptual attribution to some individual guides reference to that individual is false. This consequence is sufficient to undermine the all-attributives model.[4] Second, as Block (ms., ch. 3) points out, the appeal to tropes has the cost of untethering the all-attributives view from the science of perception. Perception science does not appeal to tropes, much less seek to explain perceptual reference to them (cp. Byrne 2019). However, we aim to assess a version of the all-attributives view that makes contact with perception science and generates testable predictions.

2.2.2—No-attributives. The no-attributives model holds that perceptual reference to objects is not guided by attributives, *punkt*. In its most extreme version, this view characterizes referential perceptual states as bare demonstratives, which simply pick out objects without attributing properties to them. However, one can simultaneously grant that perceptual states invariably attribute properties and also deny that attribution plays a role in guiding perceptual reference (Millar forthcoming).

Pylyshyn argues that the referring elements of PORs ("visual indexes") are akin to bare demonstratives. According to Pylyshyn, when PORs pick out objects, "the visual system need not know (i.e., need not have detected or encoded) any of their properties in order to implicitly treat

[3] Burge occasionally suggests that our perceptual systems secure reference to property instances as well as objects (2009, 257; 2010a, 496; cp. Schellenberg 2018), though he does not argue for it. We don't take a stand here on a separate but related issue about perceptual acquaintance with general properties (Batty & Macpherson 2016; Alford-Duguid 2018; Ivanov 2019; Millar forthcoming).

[4] Burge's case for the non-propositional That F model relies on the claim that all perceptual attribution functions to guide reference. This claim is false if there are perceptual attributions of features to objects that do not guide reference to those objects. If an attributive makes a non-reference-guiding attribution to a feature-bearer, then, by Burge's own lights, it plays a purely predicative role in the POR for that feature-bearer (even if it also guides reference in some other context).

them as though they were distinct and enduring visual tokens" (2009, 267). The idea here is not that there are no environmental properties that trigger perceptual reference, but rather that these properties need not be attributed to objects for reference to occur. The locking of reference to an object occurs "through a brute causal mechanism" (Pylyshyn 2009, 269), such that the properties to which the mechanism is sensitive need not "be *encoded* and entered into an object file in order for an index to be captured or an object file to be created" (Pylyshyn 2007, 81; emphasis his; cf. Green 2017).

While Pylyshyn often emphasizes the lack of attributional guidance in reference locking, he makes similar claims about reference maintenance. In tracking, "the identity of these objects *qua individuals* must be maintained or tracked despite changes in the individual's properties," and "this is done primitively without identifying the object through a unique descriptor" (2009, 267). Even attributives concerning the location of the object are, for Pylyshyn, not required to maintain reference during tracking. He claims, for instance, that "data on MOT suggest that the tracking system either does not have location (or direction) information or, if it does, it does not (or cannot) use it with any precision in tracking" (2007, 84). While an object's spatiotemporal, shape, and color properties are potentially causally relevant to determining perceptual reference to the object, they are not causally relevant by virtue of being encoded in the POR for that object.

In defending the no-attributives model, Pylyshyn relies largely on evidence showing that tracking succeeds despite changes or loss of surface-featural attributives. In an important unpublished manuscript, Scholl et al. (ms) used a tracking task in which objects frequently changed shape or color when temporarily invisible behind occluders. Subjects were asked to indicate whether particular objects had changed, and were no better at change detection for tracked than untracked items. Bahrami (2003) ran a similar experiment using "mud-splashes" to occlude objects during color or shape changes. Contrary to Scholl et al.'s null effect, Bahrami found superior change detection for tracked vs. untracked items; however, there were still a significant portion of trials where subjects failed to notice featural changes despite successful tracking (see also Saiki 2003; Pylyshyn 2004). These results indicate that, while deploying a POR in tracking may facilitate encoding surface properties, these properties are not invariably encoded and therefore may fail to guide reference.

Much of the tracking literature suggests that "featural information" doesn't guide reference maintenance. Here, 'features' standardly refers to surface properties like shape, color, and size, and excludes spatiotemporal properties like location and motion trajectory. Some tracking evidence suggests that spatiotemporal attributives *are* typically present in tracking (Scholl et al. ms). Keane and Pylyshyn (2006) found that when objects briefly disappeared during tracking, they were tracked more accurately when they reappeared at the location where they disappeared than when they reappeared somewhere else (even at the location one would expect given their trajectory), suggesting that a record of their last known locations was preserved and used to maintain reference to them after they reappeared. The importance of spatiotemporal properties for tracking raises the possibility that they play a special role in guiding reference, which leads us to the privileged-attributives model.

2.2.3—Privileged-attributives. According to the privileged-attributives model, some fixed class of attributives guides reference, while other attributives are (at least typically) referentially inert. The most influential form of this model lies in what Flombaum et al. call the "principle of spatiotemporal priority" (2009, 135), viz., that spatiotemporal attributives, especially location and motion trajectory, are privileged in guiding reference maintenance.[5] Flombaum et al. characterize PORs as akin to file folders (cp. Recanati 2012; Huemer et al. 2018; Murez et al. 2020), with referential components functioning as labels and non-spatiotemporal attributives as file contents:

> Spatiotemporal features can update a folder's label, perhaps even occasioning the creation of a new folder or the destruction of a current folder. Surface features, in contrast, are just along for the ride: when they change, the only result may be that the description stored inside the folder is modified.
>
> (Flombaum et al. 2009, 153)

Other proponents of the spatiotemporal priority hypothesis include Kahneman et al. (1992), Scholl (2007), and Mitroff and Alvarez (2007).

As mentioned earlier, tracking data suggests that spatiotemporal attributives are more frequently and stably encoded in PORs than other attributives. Location and trajectory also seem to play a central role in reference maintenance. For example, recall that changes in surface features like color are more likely to be detected in tracked vs. untracked items. When such changes occur while an object is occluded (e.g., a blue object goes behind an occluder and comes out red on the other side), subjects are more accurate when the object obeys spatiotemporal continuity—specifically, when the object emerges at the expected time and location given its pre-occlusion trajectory. This suggests that the object-based change detection advantage arises from spatiotemporally guided reference maintenance (Flombaum & Scholl 2006).

Other paradigms support similar conclusions. In apparent motion paradigms, subjects see displays of static objects that create an illusory visual impression of object motion. For example, if you see the left panel of Figure 1 followed by the right panel, your visual experience will represent the top and bottom objects as moving right and left, respectively. Apparent motion will follow the shortest motion path, even if surface features suggest otherwise; if the top object is a square and the bottom is a circle, for instance, then even if the second display shows a circle on top and a square on

[5] Another form of the privileged-attributives view takes categorical identity to play a privileged role in guiding reference (Gordon & Irwin 2000).

bottom (suggesting vertical motion paths), the visual system will persist in seeing horizontal motion and will simply present the objects as shifting shapes (Navon 1976; Dawson 1991). While motion perception need not be object-based, Odic et al. (2012) showed that the object-specific preview benefit follows apparent motion in these cases. The developmental literature also suggests that infants under a year old privilege spatiotemporal information in computing object correspondence even when relevant kind/featural information is successfully encoded (Xu & Carey 1996; Carey 2009, 77-83).

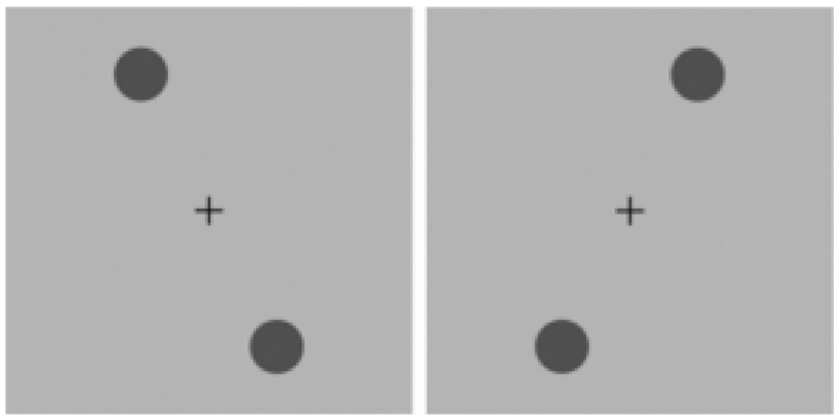

Figure 1—Apparent motion (adapted from Odic et al. 2012, Figure 2). Reprinted with permission from Taylor & Francis.

2.3—Moving forward. As we've seen in this section, there is no consensus about the reference-attribution relation in PORs. Some evidence points toward spatiotemporal priority, but as we'll now argue, other recent evidence reveals a reference-guiding role for surface-featural attributives. Thus we need a model that accommodates the guidance of reference maintenance by features such as color in some contexts and the ignoring of those same features in other contexts. The all-attributives, no-attributives, and privileged-attributives lack the flexibility required to accommodate this evidence. We will therefore develop a *flexible-attributives* model of the guidance relation between perceptual attribution and perceptual reference, according to which the class of attributives used to guide reference shifts dynamically depending on multiple factors, including (a) which properties the visual system deems most informative about object identity in a given context and (b) the current contents of attention and visual working memory.

3. A Flexible-Attributives Model

This section defends the flexible-attributives model in two stages. First we argue, *contra* the no-attributives and privileged-attributives models, that surface-featural attributives can guide reference maintenance even when spatiotemporal information is unambiguous. We then argue, *contra* the all-attributives model, that the class of attributives consulted in reference maintenance changes dynamically according to context.

3.1—Surface Features in Reference Maintenance. Object perception is studied through a multiplicity of experimental paradigms. As one would expect, evidence for the role of attributives in reference maintenance comes from a correspondingly diverse array of sources. We will now run through evidence for referential guidance by surface-featural attributives as concisely as possible.

We note first that even studies whose findings generally support the spatiotemporal priority principle reveal exceptions to the rule that surface features cannot guide reference maintenance. For example, topological properties guide object correspondence in multiple-object tracking even when other features do not. Zhou et al. (2010) ran a multiple-object tracking task in which surface properties changed unpredictably as objects moved. These surface properties included color, luminous flux, and various shape properties. They found that only topological shape changes disrupted tracking—for example, punching a hole in a circle or changing a 5 to a 9 disrupts tracking whereas changing a circle to a square or a square to an elongated rectangle does not (cp. Kibbe & Leslie 2016). A sophisticated spatiotemporal-priority hypothesis could accommodate this possibility, however, by allowing that reference-guiding spatiotemporal attributives include not only location and trajectory but also abstract spatial features including topological shape class (cp. Burge 2010, 40; Flombaum et al. 2009, 157n1).

Object-reviewing studies using the object-specific preview benefit (OSPB) provide convincing evidence that there are contexts in which processes subserving reference maintenance for PORs exploit surface features beyond topological shape class. This evidence raises problems for both the no-attributives and privileged-attributives models. Hollingworth and Franconeri (2009), for instance, found that when spatiotemporal information was insufficient for determining object correspondence, color continuity alone was sufficient to produce an OSPB. Other object-reviewing studies indicate that sudden visible *changes* in surface features can *break* object continuity. Moore et al. (2010) found that when objects abruptly changed color as they moved, the OSPB was eliminated (see also Fiedler & Moore 2015). Jiang (2020) similarly found that abrupt color changes while objects were visible significantly reduced the OSPB (as did abrupt changes in color, shape, and topology all at once).

Thus, object-reviewing studies suggest that the visual system: (i) takes continuity of color to support object continuity, at least when spatiotemporal information is ambiguous, and (ii) takes changes in color to count against object continuity. Feature (ii) is particularly noteworthy because it suggests that surface features can inform reference maintenance even when spatiotemporal information is *unambiguous*. Sudden changes in color disrupt object correspondence even if the object follows a fully visible, continuous motion trajectory. So even when spatiotemporal attributives clearly support a correspondence solution, surface-featural attributives may veto that solution, leading the visual system to instead treat the initial object as going out of existence.

Further evidence for the role of surface features in reference maintenance derives from studies of object correspondence across saccadic eye movements. Hollingworth et al. (2008) had subjects fixate the center of a circular array of colored disks. One of the disks was cued and subjects saccaded to that object. During the saccade, the array sometimes rotated slightly clockwise or counterclockwise, so the saccade landed halfway between the saccade target and a distractor. These rotations were invisible to the subjects, since visual input is suppressed during a saccade. Hollingworth et al. reasoned that if color can be used to compute object correspondence through saccades, then subjects should reliably execute a corrective saccade to the target disk rather than a different-colored distractor, even though spatial information did not favor one over the other. Consistent with this, subjects' initial corrective saccade reached the target on 98% of trials. Thus, color attributives are plausibly used to sustain reference through saccades when spatiotemporal information is ambiguous, and similar effects are found even when the object remains stationary during saccades (Poth & Schneider 2016).

While classic experiments on apparent motion suggested that surface features are ignored in computing object correspondence, more recent studies indicate otherwise. The Ternus display is an apparent motion stimulus in which a three-object array is shown first; then, after a blank period of variable length, a second frame appears in which another array is shown, shifted slightly rightward (Fig. 2). Two motion percepts are possible: *element motion*, wherein the leftmost object appears to hop across the other two objects, or *group motion*, wherein all three objects appear to move rightward. The representations underlying motion perception in Ternus displays are object-based (Stepper et al. 2019b), suggesting that object correspondence processes exploiting information contained in PORs are responsible for these effects.

Importantly, it is possible to use surface features to bias perception of the display toward either group or element motion (Fig. 2). For example, if the center object of the first frame is unique in size while the leftmost object is unique in the second frame, then (if surface features are used in correspondence) the perceiver should perceive element motion, since size continuity should lead the visual system to treat the leftmost object of the second frame as a continuation of the center object of the first frame. Conversely, if the center element is unique in both frames, this should promote perception of group motion. Hein and Moore (2012; 2014) found that perceived Ternus motion is biased in precisely this sort of way by color, size, luminance, and grating orientation. For example, when luminance induced an element-motion bias, group motion was perceived only 22% of the time, while it was perceived 85% of the time when group-motion bias was induced.[6] Further evidence demonstrates that the properties used to guide correspondence in the Ternus display are distal (Hein & Moore 2014), ruling out the view that the correspondence computations are defined over proximal features of the retinal image or inner neural layout (cf. Pylyshyn 2007, 83-85).

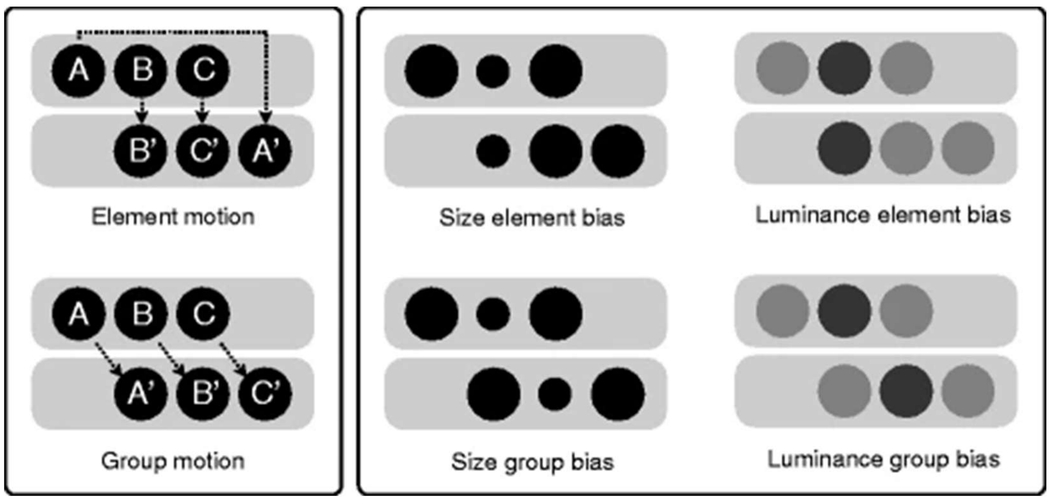

Figure 2—Ternus display from Hein & Moore (2014). Reprinted with permission from Springer.

We have discussed only a small sample of the evidence that perceptual computations of object correspondence consult surface-featural attributives. There is also evidence that diachronic continuity or similarity in surface features influences the perception of ambiguous bouncing/streaming displays (Feldman & Tremoulet 2006; Caplovitz et al. 2011), perception of launching versus passing in Michotte-style causal perception displays (Moore et al. 2020), the flash-lag effect (Moore & Enns 2004; Moore et al. 2007), and inhibition of return (Tas et al. 2012). Still more evidence will be adduced below. However, the data examined thus far is already sufficient to undermine both the no-attributives and privileged-attributives models. Attributives are consulted in reference maintenance, and these minimally include attributives for color, lightness, orientation, and size—not just spatiotemporal properties and topology.

3.2 Flexible Use of Features—We now argue for a *flexible-attributives view*, on which the range of attributives consulted in reference maintenance changes in a systematic, adaptive way depending on the perceptual context. We emphasize two types of flexibility. First, the visual system is more likely to rely on an attributive in contexts where continuity in the attributive is a reliable cue to object identity and more likely to ignore the attributive when it is an unreliable cue. Second, the visual

[6] Why does surface-featural information bias perception of the Ternus display but not other apparent motion displays? Hein and Cavanagh (2012) suggest the explanation may lie in the distance over which motion is computed. Ternus motion typically involves a small spatial range, while other stimuli involve longer inter-object distances between frames.

system is more likely to rely on attributives for a given feature when that feature (or the object that bears it) is selectively attended or retained in visual working memory. We further contend that these facts support the flexible-attributives view over the all-attributives view.

If the context indicates that one's representation of a feature is unreliable, it makes sense to rely less on that representation when reidentifying objects over time. Suppose I see a notebook on my desk and the notebook appears green. I want to know whether it's the same green notebook I left there yesterday. If I know I am in a context where my color vision is error-prone, then the notebook's apparent color provides far weaker evidence that this is the case. Likewise, facts about the object itself are relevant—color is a more reliable guide to object continuity for notebooks than chameleons. More generally, it is inadvisable to rely on a represented feature to determine object correspondence if you have reason to suspect that the feature is unstable or its visual representation is unreliable. Might the visual system employ a similar strategy in maintaining reference?

We noted earlier that surface features are often irrelevant to tracking, and hence reference maintenance. Tracking accuracy is largely unimpaired by surface feature changes, and largely unaided by surface feature stability (although see Makovski & Jiang 2009; Howe & Holcombe 2012). However, in an important modification, Papenmeier et al. (2014) examined whether surface features might play an increased role in tracking when spatiotemporal information is unreliable. They introduced brief disruptions in the middle of the task where either the entire tracking display was rotated or zoomed, or the objects "jumped" discontinuously from one location to another. When the targets all had unique colors that were distinct from the distractors, and the targets' colors remained stable during the spatiotemporal disruption, tracking performance was superior relative to both a homogenous condition, where all objects had the same color, and a swap condition, where targets swapped their colors with distractors during the disruption. Importantly, Papenmeier et al. found that swapping colors between targets and distractors did *not* impair tracking accuracy if the swaps didn't coincide with a spatiotemporal disruption. These results suggest that the visual system consulted color attributives to guide reference maintenance *only* during periods where spatiotemporal attributives were deemed unreliable.[7]

A study by Gordon et al. (2008) provides additional evidence that the visual system adopts a strategy of determining reference maintenance in accordance with those attributives that are deemed most reliable in context, but ignoring attributives deemed unreliable. Subjects who were fixating in the center of a display saw an object outside fixation (e.g., a fish facing leftward) and moved their eyes toward it; the task was to name the object after the saccade. During the saccade, one of three things happened: the object either remained at its location (same-object), an object of the same kind

[7] Likewise, a given color change is less likely to disrupt object correspondence when it occurs in the context of a series of random color changes (Au & Watanabe 2013), suggesting that the visual system also takes into account featural stability when deciding which attributives to consult in reference maintenance.

appeared in a different location (different-object), or a novel kind of object appeared (novel-kind) (Fig. 3). The object may or may not have switched its orientation (e.g., from leftward-facing to rightward-facing). In keeping with object reviewing studies, there was an object-specific preview benefit in reaction time for naming the object (e.g., “fish”) in the same-object condition. This preview benefit was diminished when orientation flipped during the saccade, suggesting that orientation was consulted to guide reference maintenance. However, when the object appeared further out toward the periphery of the scene, this effect disappeared; change in orientation no longer affected the object-specific preview benefit.

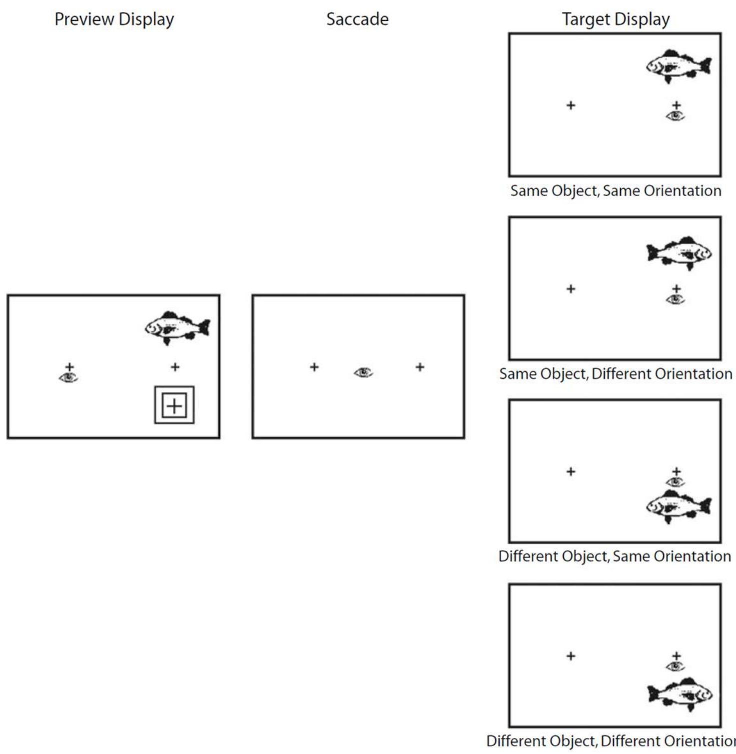

Figure 3—From Gordon et al. (2008) (novel-kind condition not depicted). Reprinted with permission from Springer.

One might worry that subjects simply failed to represent or remember orientation when the object was seen peripherally. But subjects were significantly above chance when asked to detect whether its orientation changed, showing that they successfully attributed orientation properties to the object and maintained them during the saccade. Thus orientation attributives were present in PORs whether objects appeared near or far from foveal vision, but only played a reference-guiding role in the former case.

This result is predicted by the flexible-attributives model. When the object appears near fixation, the visual system uses orientation to guide reference. But when the object appears further from fixation, the visual system determines that the poorer quality of peripheral vision means that the orientation attributive should not be relied on to guide reference, and therefore ignores it for the purposes of reference maintenance.

Finally, we mentioned above that the object-specific preview benefit is diminished by abrupt color changes during visible object motion (Jiang 2020). However, Jiang found that this effect disappears if the color change happens while objects are occluded, consistent with prior experiments on the tunnel effect (Flombaum & Scholl 2006). The flexible-attributives view can explain this discrepancy as follows. The visual system deems surface-featural attributives to have lower reliability during occlusion, so is more likely to ignore them in maintaining reference after the object reappears. Occluded objects may undergo changes for unseen reasons, while a visible object is unlikely to abruptly change its color. The visual system appears to use principles like this in determining which attributives to use in maintaining visual reference.

We now turn to the second form of flexibility, which involves changes in the visual system's reference-maintenance strategy depending on the contents of attention and working memory.

There is strong evidence that the operations responsible for solving the correspondence problem are sensitive to attention. For instance, when subjects view an ambiguous apparent motion stimulus with two equally likely interpretations, shifts between the interpretations reliably track shifts in spatial attention (Xu et al. 2013). More important for present purposes, however, recent evidence suggests that attention influences which *surface-featural attributives* the visual system relies on in maintaining reference.

We've seen that surface-featural attributives can guide reference maintenance in the Ternus display. Recently, Stepper et al. (2019b) investigated whether attention modulates this effect. In what they call the *competitive* Ternus display, some features support an element-motion bias while others support a group-motion bias. An example is shown in Figure 4, right column. Because the center elements of the first and second frame match in color, this supports group motion. However, there is also a color match between the leftmost element of the first frame and the rightmost element of the second frame, which supports element motion. Stepper et al. found that when subjects' attention was drawn to the leftmost item, element motion was significantly more likely than when attention was drawn elsewhere. Conversely, when subjects attended to the center element, group motion predominated. Thus, when an element was attended, the attributive for its color was more likely to be used in reference maintenance, while the unattended element's feature was more likely to be ignored.

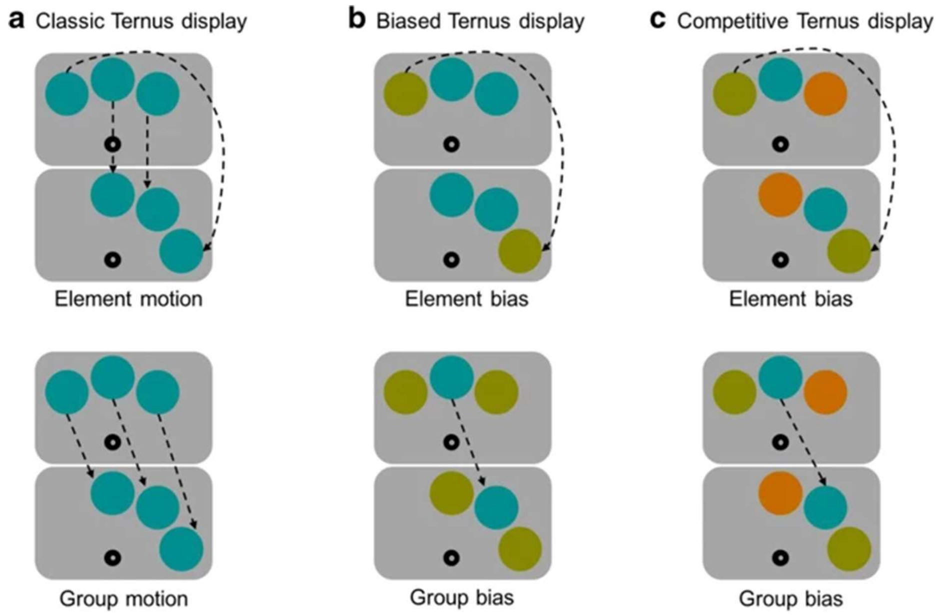

Figure 4—From Stepper et al. (2019b). Reprinted with permission from Springer.

Research on object correspondence through saccades provides evidence for the role of visual working memory in selecting which attributives to consult in reference maintenance. Richard et al. (2008) had subjects fixate the center of a display and cued them to saccade to objects in the periphery (Fig. 5). The cue indicated either the target's color or its position. On some trials, the objects shifted during the saccade so subjects had to issue a corrective saccade to the target object. In half of these shift trials, objects swapped colors as well. If a color (e.g., red) was cued, subjects often issued corrective saccades to the *currently* red object even if the colors had swapped, such that the originally cued object was no longer red—an effect that persisted even when subjects were explicitly instructed to saccade to the *position* of the cued object and ignore its post-saccade color. This result suggests that the color cue biased reference-maintenance processes to use color attributives, regardless of explicit task instructions. Richard et al. (2008, 84) explain this by appeal to visual working memory. When the cue specified the target by its color, this led its color to be retained in visual working memory, and when color was retained in memory, it was more likely to guide reference maintenance through the saccade.[8]

[8] See Hein et al. (2021) for evidence that features retained in visual working memory are also more likely to guide reference maintenance in the Ternus display.

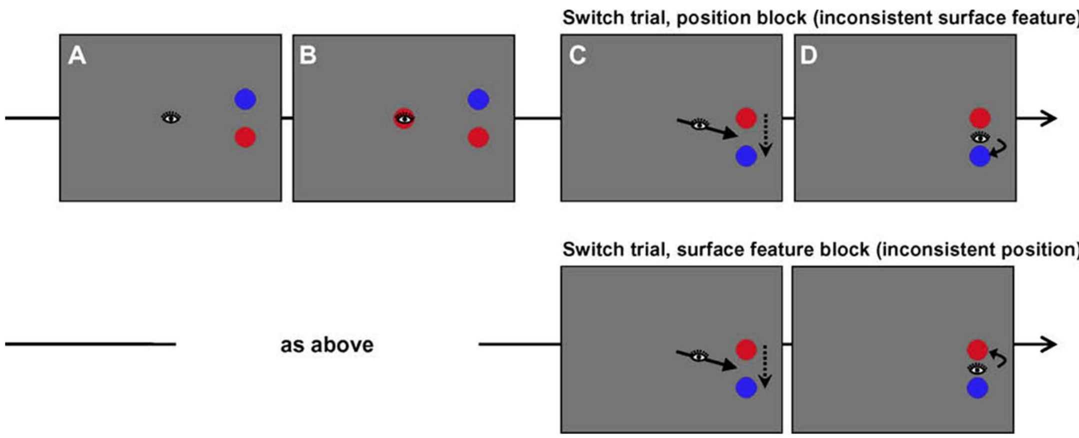

Figure 5—From Richard et al. (2008). This depicts a trial where the cue specifies the saccade target by its color. Reprinted with permission from Elsevier.

Thus, there is evidence that an attributive is more likely to be consulted in reference maintenance when it (or the feature it represents) is either attended or retained in working memory. We note that this form of flexibility in reference maintenance is probably advantageous. If it is in your interest for a feature to be prioritized in tracking (e.g., when keeping track of a yellow taxi amidst a throng of differently colored vehicles), you are more likely to attend to that feature or hold it in working memory. In turn, the visual system is more likely, other things being equal, to use that feature to guide reference.

Contra the all-attributives model, we take this evidence to suggest that not all attributive elements of a POR serve to guide reference for that POR. The visual system selectively uses different features to guide reference depending on the context. In particular, surface-featural information is consulted when it is particularly informative for object individuation or particularly salient, but ignored when it is uninformative or non-salient.

We should note that Burge (2010b, 40-41) explicitly claims only that attributives *function* to guide reference, which allows for cases of malfunction where attributives fail to guide reference but still have the function of doing so. However, the cases where perceptual attributives fail to guide reference do not seem like lapses or malfunctions. They are not pathological, or even atypical. For instance, cuing spatial position obviates the reference-guiding role of color attributives (Richard et al. 2008), but it does not do so by causing the mechanisms responsible for reference maintenance to malfunction. A more plausible view is that, in general, when an attributive is ignored in reference maintenance, the attributive neither guides nor functions to guide perceptual reference. If an attributive is not useful for locking or maintaining reference, then ignoring it does not constitute any "lapse of normal representational operation" (Burge 2010a, 311) akin to illusion or reference failure.

An anonymous reviewer notes that sometimes a whole class of items can have a function even though only some of those items actually perform it. All sperm arguably have the biological function of inseminating eggs even though few of them actually do so.[9] Likewise, it might be suggested that on any given occasion, all perceptual attributives function to guide reference even though only some of them actually do so.

However, the considerations that favor positing an infrequently performed function in the sperm case do not carry over to perceptual attributives. What makes it the case that sperm have the rarely fulfilled function of inseminating eggs rather than, say, a more frequently fulfilled function like entering the cervix is, plausibly, that sperm that enter the cervix only promote the reproduction of their producing system if they also inseminate an egg (e.g., Millikan 1989; Neander 2017, ch. 3). However, no parallel argument is available in the case of perceptual attributives. Suppose we are right that not all perceptual attributives actually contribute to reference guidance. For the sperm analogy to work, one would need to show that among those attributives within a POR, only those that also guide reference for the POR are efficacious in producing some adaptive benefit for the system. But there is little reason to believe this. If I see a ripe apple and my perceptual system accurately represents its color via the attributive *red*, then I am in a position to categorize the object as an apple and form true beliefs about its color and ripeness, which may in turn lead me to eat it and reap its nutrients. Nothing in this story requires that *red* also guide perceptual reference to the apple. Perhaps the apple was singled out wholly by its shape and location. There is not even a good reason to think that the outcome would have been *more* adaptive had *red* played a reference-guiding role. Thus, unlike sperm inseminating eggs, perceptual attributives needn't guide reference in order to produce adaptive benefits.[10]

The flexible-attributives model posits an intelligent perceptual system that has access to information about the visual context, such as the reliability of various attributives at the moment. If spatial trajectory is unclear, for example, then reference-maintenance processes might instead use color as an indication of which object went where; but in other contexts trajectory or other spatiotemporal properties may dominate, leading color attributives to be ignored.

The flexible-attributives model permits there to be contexts where a property is attributed to an object without being used to guide reference to that object. However, one might suggest an alternative flexible model more congenial to the all-attributives view. This model would hold that all

[9] We will assume for the sake of argument that all sperm have this function, although this may not be true on more demanding theories such as Shea's (2018, ch. 3) notion of robust task function.

[10] We note that Burge (2010b) seems to regard the reference-guiding function of perceptual attributives as a *representational* function rather than a biological function. Representational functions are, roughly, functions associated with veridicality (and perhaps reference), rather than biological fitness (Burge 2010a, 308-310). Perceptual systems fulfill their representational functions when they represent veridically, and fail to fulfill them in cases of illusion or reference failure. In any case, if an attributive is unreliable for purposes of reference maintenance, then ignoring it in reference maintenance doesn't seem like any failure of representational function. Rather, if an attributive indicates a feature that is an unreliable cue to object continuity in context, then ignoring that attributive in reference guidance is *conducive* to fulfilling the representational function of sustaining perceptual reference to a persisting individual over time.

attributives are consulted in reference maintenance, but the importance assigned to them can be weighted up or down if they are deemed more or less useful or salient in a particular context (cp. Feldman & Tremoulet 2006; Hein & Moore 2012; Papenmeier et al. 2014). For instance, if color is weighted higher, then the visual system is more likely to take color continuity to support object continuity, and color discontinuity is more likely to break object continuity. This weighting-based flexible model could be construed as a version of the all-attributives model, albeit one on which some attributives can be overruled by others in reference determination.

However, in many of the cases discussed above, attributives seem to have no effect *at all* on reference maintenance. For example, surface-featural changes that are encoded in PORs (Bahrami 2003) fail to disrupt tracking performance (Zhou et al. 2010). Furthermore, some measures of reference maintenance are graded, allowing greater sensitivity to smaller effects than binary measures like tracking (successful vs. unsuccessful) or the Ternus display (group vs. element motion).

For example, on trials where the saccade targets moved in Richard et al.'s (2008) study, subjects had to re-identify the target and issue a "corrective saccade" to its new location. One of their dependent variables was the duration of gaze prior to the corrective saccade, which provides a temporal index of the reference-maintenance processes required to re-identify the target. If color, for example, is invariably used in reference maintenance (Papenmeier et al. 2014), then a color change should increase gaze duration in all contexts. For if color attributives are obligatorily consulted alongside spatiotemporal attributives, then we would expect the reference-maintenance computations required for saccade correction to take longer when color and position cues to correspondence disagree, since the perceptual system needs to resolve the conflict. However, while Richard et al. found that context can affect which features are "preferentially weighted," they also found that when position was effectively cued, the effect of surface features like color was not merely reduced but "eliminated entirely" (2008, 82). Thus: "[I]t appears that participants can limit transsaccadic memory and/or comparison operations to position information when the task is configured so that the target is selected solely on the basis of position" (2008, 82).

Similarly, the object-specific preview benefit involves a quantitative change in reaction time that could be greater or lesser rather than simply appearing or not. Thus changes in one attributive will reduce the preview benefit, but not as much as changes in multiple attributives. Jiang found that the preview benefit was "reduced but not eliminated" given unoccluded color changes (2020, 465), while unoccluded changes on multiple dimensions at once eliminated it altogether. When an object changed its features while occluded, however, the preview benefit was not significantly affected. For objects that disappeared behind occluders and re-emerged, if they re-emerged a different color the preview benefit was 57ms, compared to 55ms when color didn't change; this difference is insignificant (and not even trending in the right direction) (Jiang 2020, 459). Similar null effects were found by Gordon et al. (2008): changing orientation during saccades diminished the preview benefit for objects close to the fovea, but had no significant effect on objects in the periphery. These results suggest that, in many contexts, such as occlusion or peripheral vision, features like color and orientation are not simply downweighted—they are wholly ignored in reference maintenance.

An anonymous reviewer suggests that Bayesian models of perception (e.g., Feldman & Tremoulet 2006) might support the idea that attributives are always consulted in reference maintenance even if they are downweighted to a degree that renders their influence unobservable, on grounds that optimal processors should never discard relevant information. We don't take a stand on Bayesian approaches to perception here (cf. Won et al. 2021). Prominent Bayesian approaches in psychology, however, often include a constraint of "resource-rationality" (Lieder & Griffiths 2019): the mind should implement algorithmic-level processes that maximize optimality while minimizing computational costs. On resource-rational Bayesian approaches, it's often desirable to run computations that fail to access potentially relevant information because a cheaper computation suffices. Following Icard and Goodman (2015), Lieder and Griffiths argue that "it is often resource-rational to ignore all but the one to three most relevant variables" in a range of computational tasks (2019, 10).

Object correspondence must be computed rapidly and dynamically, generating non-trivial costs to adding additional variables to the computation. There is a cost not only in computing over the attributive, but also assigning it a precise weight and balancing it against other attributives involved in the correspondence computation. If there are contexts where the benefits of determining an appropriate weight and computing over an attributive are negligible and outweighed by the costs of doing so, it would be resource-rational for the visual system to identify those contexts and discount the relevant attributives altogether. Just this sort of context may be realized when an object's features have potentially changed due to unobservable events (e.g., during occlusion, during long saccades, or when the precision of the attributive is low to the point of unreliability), rendering the attributive more cost than it's worth for reference maintenance. If the object is in full, high-resolution view and it's unlikely to have rotated, then consulting its orientation is a useful strategy for reference maintenance; but if it's gone unobserved for some time or the representational quality is low, then plausibly it's more useful to ignore certain surface features and rely on (e.g.) location and trajectory instead.

Thus, there is empirical evidence against the downweighting model, and considerations of optimality do not clearly support the downweighting model over the flexible-attributives model. We also note that there is no reason to think that attributives themselves, and the information they contain, are discarded in the studies we've mentioned. Instead, they simply fail to be consulted in a particular computation (object correspondence), remaining available for other processes such as change detection (Gordon et al. 2008).[11]

[11] In saying that certain attributives are ignored in certain perceptual computations, we also don't mean to suggest a general model of cue integration where one sensory cue to a feature dominates while the others exert no influence on the resulting percept. While such models were once common, particularly in early research on multisensory interaction (e.g.,

An important consequence of the flexible-attributives model is that the role attributives play in guiding reference is dynamic. While different feature dimensions might be organized into different slots within a POR (Green & Quilty-Dunn 2017), the visual system does not impose context-independent restrictions on which of these slots get to play a reference-guiding role. Among the various properties attributed to an object, some can guide reference in some contexts, but in other contexts the very same property-attributions play no role at all in guiding reference. The visual system thus seems not only to flexibly weight attributives in reference maintenance, but also to make a binary decision about whether a particular attributive should be consulted for reference-maintenance in a particular context.[12] The models discussed in Section 2 have trouble accommodating the flexibility of reference guidance, and the view that all features are weighted to varying degrees has trouble explaining cases where attributives are present but play no guiding role whatsoever. The flexible-attributives model—which allows attributives to be flexibly weighted, but also posits a binary distinction between attributives that are and aren't consulted for reference guidance in a context—is explanatorily superior.

Before concluding, we consider one more reply on behalf of the all-attributives view. Earlier we distinguished between two forms of reference determination: reference *locking* and reference *maintenance*. So far, we have argued that not all attributives are consulted in reference maintenance. However, a proponent of the all-attributives view might insist that every perceptual attributive serves to guide *either* reference locking *or* reference maintenance. If so, then even if an attributive fails to guide reference maintenance in some context, this is not a problem as long as it was at least involved in locking (i.e., initially selecting the relevant particular).[13]

Rock & Victor 1964), they have been superseded by models wherein the perceptual system combines various cues to a feature in a manner weighted by their relative reliability (inverse variance) (van Dam et al. 2014; Rescorla 2018). Our view is compatible with these modern conceptions of cue integration. Specifically, we do not deny that when attributives are consulted in a perceptual computation, they are all assigned some weight rather than one simply dominating the others. But we claim that there are cases where certain attributives just aren't consulted at all, and thus do not factor into the cue integration computation. Resource-rational accounts offer theoretical motivation for this sort of framework, suggesting that it may best meet the demands of optimality.

[12] Notice that this model leaves room for a weakened version of the spatiotemporal priority principle. It may be that while it is possible to wholly ignore surface features in reference maintenance (viz., if they are excluded at the binary decision stage), it is not possible to wholly ignore spatiotemporal attributives (cp. Richard et al. 2008, 82). Perhaps spatiotemporal attributives must always be assigned some weight in reference maintenance, even if they don't invariably dominate surface-featural attributives. Furthermore, Xu & Carey's (1996) evidence that object correspondence in 10-month-olds relies on spatiotemporal features and *cannot* incorporate featural/kind information until later in development suggests that spatiotemporal information may also have a privileged role in the development of object perception. Perhaps the reference-guiding role for spatiotemporal attributives comes "online" first, and only later does the visual system come to use other information to guide perceptual reference.

[13] Relatedly, one might suggest that when attributives aren't used for reference maintenance, they still aid in segregating an object from its surroundings. However, our present arguments concern the reference-guiding role of attributions *in*

While this is a tempting option, we contend that if the arguments we've advanced so far are successful, then this view is untenable. Once one concedes that certain attributives fail to guide reference maintenance, it is a short step to the conclusion that certain attributives guide neither reference maintenance nor reference locking.

The strongest evidence that not all attributives are used in reference maintenance involves cases where a change in some feature dimension is perceptually represented, but the processes responsible for reference maintenance treat the change as irrelevant to the continuity of the object that undergoes it. Suppose, then, that in a given context the visual system deems color attributives unreliable for reference maintenance, so they are ignored. Nevertheless, color is still attributed to the object in question. Suppose further that the object visibly changes from red at time T1 to blue at T2, and although it is tracked throughout, neither the attributive *red* nor the attributive *blue* is used to maintain reference to it. Now consider the attribution of blueness at T2. We've already said that this attribution plays no role in reference maintenance. Crucially, however, it cannot assist in reference *locking* either, because the attribution of blueness occurs after reference was already secured to the object. Thus, if a feature dimension is ignored in reference maintenance, then it is almost certainly true that some attributions of features along that dimension (viz., attributions that occur after reference locking is complete) guide neither reference maintenance nor reference locking.

4. Conclusion

When you perceive a red tomato, your perceptual state both refers to a particular object (the tomato) and attributes properties to it (redness and roundness). This paper has examined the role of perceptual attribution in determining perceptual reference. We considered three models of this role and argued that none of them succeed. Perceptual attributives are used to guide perceptual reference (*contra* the no-attributives model), and the range of attributives so used is not limited to spatiotemporal and topological attributives (*contra* the privileged-attributives model). Nevertheless, within a context the visual system often elects to ignore attributives that it deems unimportant or unreliable for purposes of reference guidance (*contra* the all-attributives model). In place of these approaches, we recommend a flexible-attributives model, on which the range of attributives consulted in reference guidance changes dynamically in systematic and adaptive ways depending on the perceptual context. This view underscores the intelligence and flexibility of our perceptual capacities. It also has implications for the broader function of perceptual attribution. While perceptual attribution *can* function to guide reference, it does not *always* function this way.

We conclude by considering some consequences of the flexible-attributives model for the *format* and *content* of perceptual representation. Regarding content, we mentioned above that

---

PORs. Figure-ground segregation occurs earlier and is less capacity-limited than the assignment of PORs (Rensink 2000; Pylyshyn 2006), and is therefore unlikely to employ the attributive elements of PORs.

Burge's claim that perceptual content is invariably non-propositional rests on an analogy with complex demonstratives, which is then cashed out in terms of the attribution-reference relation: perceptual attribution has a *That F* form because perceptual attributives always function to guide reference. Perceptual attributives therefore never occur outside the scope of a noun phrase, and so cannot suffice for "pure predication" as exhibited by the predicate *G* in *That F is G* (Burge 2009). Since pure predication is a prerequisite for propositional structure, it follows that perceptual content is never propositional. This difference is claimed to ground a categorical distinction between the content of perception and the content of thought (Burge 2010b).

Reasonable doubts can be raised about various steps of this argument, such as: the claim that propositional content requires pure predication; the claim that pure predication cannot play the metasemantic role of guiding reference; the claim that nonpropositional content in perception is accurately modeled by complex demonstratives. But our objection in this paper is a more direct empirical challenge: it is not true that perceptual attributives invariably play a reference-guiding role, so Burge's (2010b) central argument for a nonpropositional model of perceptual content is unsuccessful.[14] Thus, the arguments given here effectively defuse Burge's characterization of the distinction between perceptual attribution and predication. However, we do not take them to *establish* that perceptual content is propositional, in part because we hesitate to endorse Burge's claim that predication is distinguished from other forms of attribution by its relation to reference. But this argument may form a piece of a larger argument for propositional content in perception (Quilty-Dunn forthcoming).

What about the format of perceptual representation? PORs are also referred to as object files, and the notion of a "file" structure pops up in many other areas of cognitive science, including work on dynamic semantics (Heim 1982), the structure of concepts (Laurence & Margolis 1999; Fodor 2008), the nature of singular thought (Jeshion 2010; Recanati 2012), and theory of mind (Perner et al. 2015; Huemer et al. 2018). It's unclear whether any natural psychological kind unites these various appeals to files (Murez et al. 2020). Perner and Leahy argue that the utility of the file framework is to "capture the predicative structure of language, i.e., the distinction between what one is talking about (the subject, topic, i.e., what the file tracks) and what one says about it (the information about the topic, i.e., the information the file has on it)" (2016, 494). They argue that mental files are used to track individuals under a particular description or *label*, which plays a different role from the *predicates* that are stored in the file's contents. We can understand the label/predicate distinction as a difference in the syntactic position occupied by various attributives, some of which are used as reference-guiding labels ('the baker', 'the ball', etc.) and others which are

[14] Burge (2010a) offers other arguments for non-propositional content in perception, hinging largely on the fact that perceptual capacities don't require capacities for rational thought. These issues overlap significantly with the debate about *nonconceptual content* (Byrne 2005; Toribio 2011). We're ignoring these issues here.

merely attributed ('is tall', 'is bouncy', etc.). From the perspective of this mental-file framework, the flexible-attributives model suggests that the format of PORs incorporates a label/predicate distinction, and also that which attributives occupy the label position is dynamically updated according to the visual context.

In sum, any viable view of the relation between perceptual attribution and perceptual reference must be responsive to the growing body of evidence that perceptual attributives play different computational roles in perceptual processing depending on context. Rigid frameworks like the no-attributives, privileged-attributives, and all-attributives models do not meet this demand, while the flexible-attributives model does. The view also has deep implications for debates about the format and content of perceptual representation.[15]

References

Alford-Duguid, D. (2020). Thinking through illusion. *European Journal of Philosophy* 28(3), 617–638.

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, 1–10.

Au, R. K., & Watanabe, K. (2013). Object motion continuity and the flash-lag effect. *Vision Research*, *92*, 19–25.

Bahrami, B. (2003). Object property encoding and change blindness in multiple object tracking. *Visual Cognition*, *10*(8), 949–963.

Batty, C., & Macpherson, F. (2016). Redefining illusion and hallucination in light of new cases. *Philosophical Issues* 26(1), 263–296.

Block, N. (ms.) *The Border Between Seeing and Thinking*.

Burge, T. (2009). Five theses on de re states and attitudes. In J. Almog & P. Leonardi (eds.), *The Philosophy of David Kaplan*, 246–316. Oxford: Oxford University Press.

Burge, T. (2010a). *Origins of Objectivity*. Oxford: Oxford University Press.

Burge, T. (2010b). Steps toward origins of propositional thought. *Disputatio*, *4*(29), 39–67.

Byrne, A. (2005). Perception and conceptual content. In E. Sosa and M. Steup (eds.), *Contemporary Debates in Epistemology* (London: Basil Blackwell), 231–250.

Byrne, A. (2019). Schellenberg's capacitism. Analysis, 79(4), 713-719.

[15] For helpful feedback, we are grateful to Nick Shea, Barry Smith, and other audience members at an Institute of Philosophy lab meeting, and to Chris Peacocke and other members of the Philosophy of Mind Work-in-Progress Discussion Group. Thanks also to Zoe Jenkin and Eric Mandelbaum for valuable discussion. We are also indebted to an anonymous reviewer for comments that led to significant improvements.

Campbell, J. (2002). *Reference and Consciousness*. Oxford: Oxford University Press.
Caplovitz, G. P., Shapiro, A. G., & Stroud, S. (2011). The maintenance and disambiguation of object representations depend upon feature contrast within and between objects. *Journal of vision*, *11*(14):1, 1–14.
Carey, S. (2009). *The Origin of Concepts*. Oxford: Oxford University Press.
Chesney, D. L., & Haladjian, H. H. (2011). Evidence for a shared mechanism used in multiple–object tracking and subitizing. *Attention, Perception, & Psychophysics*, *73*(8), 2457–2480.
Clark, A. (2004). Feature-placing and proto-objects. *Philosophical Psychology*, *17*(4), 443–469.
Cohen, J. (2004). Objects, places, and perception. *Philosophical Psychology*, *17*(4), 471–495.
Dawson, M. R. (1991). The how and why of what went where in apparent motion: modeling solutions to the motion correspondence problem. *Psychological Review*, *98*(4), 569–603.
Feldman, J., & Tremoulet, P. D. (2006). Individuation of visual objects over time. *Cognition*, *99*(2), 131–165.
Fiedler, A., & Moore, C. M. (2015). Illumination frame of reference in the object-reviewing paradigm: A case of luminance and lightness. *Journal of Experimental Psychology: Human Perception and Performance*, *41*(6), 1709–1717.
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, 840–853.
Flombaum, J.L., Scholl, B.J., & Pylyshyn, Z.W. (2008). Attentional resources in visual tracking through occlusion: The high-beams effect. *Cognition* 107(3), 904–931.
Flombaum, J.I., Scholl, B.J., & Santos, L.R. (2009). Spatiotemporal priority as a fundamental principle of object persistence. In B. M Hood & L.R. Santos (eds.), *The Origins of Object Knowledge*, 135–164. Oxford: Oxford University Press.
Fodor, J.A. (2007). The revenge of the given. In B. McLaughlin and J. Cohen (eds.), *Contemporary Debates in Philosophy of Mind* (Oxford: Blackwell), 105–116.
Fodor, J.A. (2008). *LOT 2: The Language of Thought Revisited*. Oxford: OUP.
Gao, T., Gao, Z., Li, J., Sun, Z., & Shen, M. (2011). The perceptual root of object-based storage: An interactive model of perception and visual working memory. *Journal of Experimental Psychology: Human Perception and Performance*, *37*(6), 1803–1823.
Gao, T., & Scholl, B.J. (2010). Are objects required for object-files? Roles of segmentation and spatiotemporal continuity in computing object persistence. *Visual Cognition* 18(1), 82–109.
Gordon R.D., Vollmer S.D., & Frankl M.L. (2008). Object continuity and the transsaccadic representation of form. *Perception and Psychophysics* 70, 667–679.
Green, E. J. (2017). Attentive visual reference. *Mind & Language*, *32*(1), 3–38.
Green, E. J. (2019). Binding and differentiation in multisensory object perception. *Synthese*. https://doi.org/10.1007/s11229–019–02351–1.

Green, E. J., & Quilty–Dunn, J. (2017). What is an object file?. *The British Journal for the Philosophy of Science*. https://doi.org/10.1093/bjps/axx055.

Haladjian, H. H., & Pylyshyn, Z. W. (2008). Object-specific preview benefit enhanced during explicit multiple object tracking. *Journal of Vision*, *8*(6), 497.

Heim, I. (1982). *The Semantics of Definite and Indefinite Noun Phrases*. PhD Dissertation, University of Massachusetts, Amherst.

Hein, E., & Cavanagh, P. (2012). Motion correspondence in the Ternus display shows feature bias in spatiotopic coordinates. *Journal of Vision*, *12*(7):16, 1–14.

Hein, E., & Moore, C. M. (2012). Spatio-temporal priority revisited: The role of feature identity and similarity for object correspondence in apparent motion. *Journal of Experimental Psychology: Human Perception and Performance*, *38*(4), 975–988.

Hein, E., & Moore, C. M. (2014). Evidence for scene-based motion correspondence. *Attention, Perception, & Psychophysics*, *76*(3), 793–804.

Hein, E., Stepper, M. Y., Hollingworth, A., & Moore, C. M. (2021). Visual working memory content influences correspondence processes. Journal of Experimental Psychology: Human Perception and Performance, 47(3), 331-343.

Hill, C. S. (2019). Perceptual existentialism sustained. *Erkenntnis*. https://doi.org/10.1007/s10670–019–00160–z.

Hollingworth, A., Richard, A.M., & Luck, S.J. (2008). Understanding the function of visual short-term memory: Transsaccadic memory, object correspondence, and gaze correction. *Journal of Experimental Psychology: General* 137(1), 163–181.

Hollingworth, A., & Franconeri, S. L. (2009). Object correspondence across brief occlusion is established on the basis of both spatiotemporal and surface feature cues. *Cognition*, *113*(2), 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), 543–564.

Howe, P.D.L., & Holcombe, A.O. (2012). The effect of visual distinctiveness on multiple object tracking performance. *Frontiers in Psychology*, *3*(307), 1–7.

Huemer, M., Perner, J., & Leahy, B. (2018). Mental files theory of mind: When do children consider agents acquainted with different object identities? *Cognition* 171, 122–129.

Icard, T. & Goodman, N.D. (2015). A resource-rational approach to the causal frame problem. Proceedings from the 37th annual meeting of the Cognitive Science Society.

Irwin, D. E. (1992). Memory for position and identity across eye movements. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*(2), 307–317.

Ivanov, I. (2021). Properties in sight and in thought. *Synthese* 198, 7049–7071.

Jeshion, R. (2010). Singular thought: Acquaintance, semantic instrumentalism, and cognitivism. In R. Jeshion (ed.), *New Essays on Singular Thought*, 105–140. Oxford: Oxford University Press.

Jiang, H. (2020). Effects of transient and nontransient changes of surface feature on object correspondence. *Perception*, *49*(4), 452–467.

Kahneman, D., Treisman, A., & Gibbs, B. J. (1992). The reviewing of object files: Object-specific integration of information. *Cognitive Psychology*, *24*(2), 175–219.

Keane, B. P., & Pylyshyn, Z. W. (2006). Is motion extrapolation employed in multiple object tracking? Tracking as a low-level, non-predictive function. *Cognitive Psychology*, *52*(4), 346–368.

Kibbe, M.M., & Leslie, A.M. (2016). The ring that does not bind: Topological class in infants' working memory for objects. *Cognitive Development* 38, 1–9.

Lande, K. (Unpublished). Seeing and visual reference.

Laurence, S. & Margolis, E. (1999). Concepts and cognitive science. In their *Concepts: Core Readings* (Cambridge, MA: MIT Press), 3–81.

Lieder, F., & Griffiths, T.L. (2019). Resource-rational analysis: Understanding human cognition as the optimal use of limited computational resources. Behavioral and Brain Sciences 43, 1–60.

Makovski, T., & Jiang, Y.V. (2009). Feature binding in attentive tracking of distinct objects. *Visual Cognition* 17(1–2), 180–194.

Meyerhoff, H. S., & Scholl, B. J. (2018). Auditory-induced bouncing is a perceptual (rather than a cognitive) phenomenon: Evidence from illusory crescents. *Cognition*, *170*, 88–94.

Millar, B. (Forthcoming). Perceiving properties versus perceiving objects. *Analytic Philosophy*.

Mitroff, S. R., & Alvarez, G. A. (2007). Space and time, not surface features, guide object persistence. *Psychonomic Bulletin & Review*, *14*(6), 1199–1204.

Moore, C. M., & Enns, J. T. (2004). Object updating and the flash-lag effect. *Psychological Science*, *15*(12), 866–871.

Moore, C. M., Mordkoff, J. T., & Enns, J. T. (2007). The path of least persistence: Object status mediates visual updating. *Vision Research*, *47*(12), 1624–1630.

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.

Moore, C. M., Stephens, T., & Hein, E. (2020). Object correspondence: Using perceived causality to infer how the visual system knows what went where. *Attention, Perception, & Psychophysics*, *82*(1), 181–192.

Murez M., Smortchkova J., & Strickland B. (2020). Mental files: an empirical perspective. In R. Goodman, J. Genone, & N. Kroll (eds.), *Mental Files and Singular Thought* (Oxford: OUP), 107–142.

Navon, D. (1976). Irrelevance of figural identity for resolving ambiguities in apparent motion. *Journal of Experimental Psychology: Human Perception and Performance*, *2*(1), 130–138.

Neander, K. (2017). *A Mark of the Mental*. Cambridge, MA: MIT Press.

Noles, N. S., Scholl, B. J., & Mitroff, S. R. (2005). The persistence of object file representations. *Perception & Psychophysics*, *67*(2), 324–334.

Odic, D., Roth, O., & Flombaum, J. I. (2012). The relationship between apparent motion and object files. *Visual Cognition*, 20(9), 1052–1081.

Papenmeier, F., Meyerhoff, H. S., Jahn, G., & Huff, M. (2014). Tracking by location and features: Object correspondence across spatiotemporal discontinuities during multiple object tracking. *Journal of Experimental Psychology: Human Perception and Performance*, *40*(1), 159–171.

Perner, J., Huemer, M., & Leahy, B. (2015). Mental files and belief: A cognitive theory of how children represent belief and its intensionality. *Cognition* 145, 77–88.

Perner, J., & Leahy, B. (2016). Mental files in development: Dual naming, false belief, identity and intensionality. *Review of Philosophy and Psychology* 7, 491–508.

Pollatsek, A., Rayner, K., & Collins, W. E. (1984). Integrating pictorial information across eye movements. *Journal of Experimental Psychology: General*, *113*(3), 426–442.

Poth, C.H., & Schneider, W.X. (2016). Breaking object correspondence across saccades impairs object recognition: The role of color and luminance. *Journal of Vision* 16, 1.

Pylyshyn, Z. W. (2003). *Seeing and Visualizing: It's Not What You Think*. Cambridge, MA: MIT Press.

Pylyshyn, Z.W. (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. (2006). Some puzzling findings in multiple object tracking (MOT): II. Inhibition of moving nontargets. *Visual Cognition*, *14*(2), 175–198.

Pylyshyn, Z. W. (2007). *Things and Places: How the Mind Connects with the World.* Cambridge, MA: MIT Press.

Pylyshyn, Z. (2009). The empirical case for bare demonstratives in vision. In C. Viger and R. J. Stainton (Eds.), *Compositionality, Context, and Semantic Values: Essays in Honour of Ernie Lepore* (Springer), 255–274.

Pylyshyn, Z. W., & Storm, R. W. (1988). Tracking multiple independent targets: Evidence for a parallel tracking mechanism. *Spatial Vision*, *3*(3), 179–197.

Quilty-Dunn, J. (2019). Perceptual pluralism. *Noûs*, doi: 10.1111/nous.12285.

Quilty-Dunn, J. (Forthcoming). Concepts and predication from perception to cognition. *Philosophical Issues*.

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

Rensink, R. A. (2000). The dynamic representation of scenes. *Visual Cognition*, *7*(1–3), 17–42.

Rescorla, M. (2018). Perceptual co-reference. *Review of Philosophy and Psychology*, doi: 10.1007/s13164–018–0411–6.

Richard, A. M., Luck, S. J., & Hollingworth, A. (2008). Establishing object correspondence across eye movements: Flexible use of spatiotemporal and surface feature information. *Cognition*, *109*(1), 66–88.

Rock, I., & Victor, J. (1964). Vision and touch: An experimentally created conflict between the two senses. Science, 143, 594-596.

Saiki, J. (2003). Feature binding in object-file representations of multiple moving items. *Journal of Vision* 3, 6–21.

Schellenberg, S. (2018). *The Unity of Perception*. Oxford: Oxford University Press.

Schellenberg, S. (2020). Capacities first. *Philosophy and Phenomenological Research*, *100*(3), 744–757.

Scholl, B. J. (2007). Object persistence in philosophy and psychology. *Mind & Language*, *22*(5), 563–591.

Scholl, B. J., & Leslie, A. (1999). Explaining the infant's object concept: Beyond the perception/cognition dichotomy. In E. Lepore & Z. W. Pylyshyn (eds.), *What Is Cognitive Science?*, 26–73. Oxford: Blackwell.

Scholl, B.J., & Pylyshyn, Z.W. (1999). Tracking multiple items through occlusion: Clues to visual objecthood. *Cognitive Psychology* 38, 259–290.

Scholl, B.J., Pylyshyn, Z.W., & Franconeri, S.L. (Unpublished). The relationship between property–encoding and object–based attention: Evidence from multiple object tracking.

Schut, M. J., Fabius, J. H., Van der Stoep, N., & Van der Stigchel, S. (2017). Object files across eye movements: Previous fixations affect the latencies of corrective saccades. *Attention, Perception, & Psychophysics*, *79*(1), 138–153.

Shea, N. (2018). *Representation in Cognitive Science*. Oxford: Oxford University Press.

Smith, A. D. (2002). *The Problem of Perception*. Cambridge: Cambridge University Press.

Smortchkova, J., & Murez, M. (forthcoming). Representational kinds. In J. Smortchkova, K. Dolega, & T. Schlicht (eds.), *What are Mental Representations?* Oxford: Oxford University Press.

Spelke, E. (1988). Where perceiving ends and thinking begins: The apprehension of objects in infancy. In A. Yonas (ed.), *Perceptual Development in Infancy: The Minnesota Symposia on Child Psychology, Vol. 20*, 197–234. Hillsdale, NJ: Lawrence Erlbaum.

Stepper, M. Y., Moore, C. M., Rolke, B., & Hein, E. (2019a). The role of object history in establishing object correspondence. *Attention, Perception, & Psychophysics*, *82*, 1038–1050.

Stepper, M. Y., Rolke, B., & Hein, E. (2019b). How voluntary spatial attention influences feature biases in object correspondence. *Attention, Perception, & Psychophysics*, *82*, 1024–1037.

Tas, A. C., Dodd, M. D., & Hollingworth, A. (2012). The role of surface feature continuity in object-based inhibition of return. *Visual Cognition*, *20*(1), 29–47.

Toribio, J. (2011). Compositionality, iconicity, and perceptual nonconceptualism. *Philosophical Psychology* 24(2), 177–193.

Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, *12*(1), 97–136.

Ullman, S. (1979). *The Interpretation of Visual Motion*. Cambridge, MA: MIT Press.

van Dam, L. C. J., Parise, C. V., & Ernst, M. O. (2014). Modeling multisensory integration. In D. J. Bennett & C. Hill (eds.), Sensory Integration and the Unity of Consciousness. Cambridge, MA: MIT Press, pp. 209-229.

Won, I., Gross, S., & Firestone, C. (2021). "Impossible" somatosensation and the (ir)rationality of perception. *Open Mind* 5, 30–41.

Xu, F., & Carey, S. (1996). Infants' metaphysics: The case of numerical identity. *Cognitive Psychology*, *30*(2), 111–153.

Xu, Y., Suzuki, S., & Franconeri, S. L. (2013). Shifting selection may control apparent motion. *Psychological Science*, *24*(7), 1368–1370.

Zhou, K., Luo, H., Zhou, T., Zhuo, Y., & Chen, L. (2010). Topological change disturbs object continuity in attentive tracking. *Proceedings of the National Academy of Sciences*, *107*(50), 21920–21924.

Zmigrod, S., Spapé, M., & Hommel, B. (2009). Intermodal event files: Integrating features across vision, audition, taction, and action. *Psychological Research*, *73*, 674–684.