

**Visual Demonstratives of Objects? Regions? Events? or Properties?
03/29/2013**

1. Abstract

To account for how we track objects, Pylyshyn (2007) and Kahneman *et al.* (1992) have postulated object representations at the early stages of cognitive processing whose referents, like those of linguistic demonstratives, are context dependent and singular. But they adopt two different accounts as to how these representations refer. Whereas Pylyshyn treats object representations as referring directly, without describing their referent in any manner, Kahneman *et al.* have the representations latching onto their referent *via* location representations. Here I first argue that Pylyshyn's adopted account of reference is inadequate because it results in events being the referents of object representations. Then, I present an account of how object representations refer in the spirit, but going beyond, of Kahneman *et al.* I argue that object representations at the early stages of visual processing refer to objects in virtue of entering in stored in memory predications with spatio-temporal predicates, where each stored predication can involve only one spatio-temporal predicate.

Key words: visual representation, visual demonstrative, early visual system, FINST, object-file, functional role, perceptual tracking, Spelke-objects, 3D and 4D accounts of persisting through time

2. Introduction

Object representations are commonly appealed to in explaining early visual processes. But there are two camps. Led by Pylyshyn (2007) one camp treats object representations as representing directly, without describing their referent in any manner. Led by Kahneman *et al.* (1992) the other camp has object representations latching onto their referent *via* location representations. The purpose of this paper is to help resolve this debate by analysing how object representations acquire their content.

How object representations in early vision refer matters not only for determining how the mind connects with the world. It is also important in determining how the object representations in early vision relate to the conceptual representations of objects¹. The latter are explicitly represented *via* descriptive concepts such as 'bounded', 'moving as a whole', and 'persisting through time'.² If Pylyshyn is right, then the early object representations can contribute merely their referent to the later, conceptual ones--they have nothing else to offer, after all. On the other hand, if Kahneman and Treisman are correct, the early representations of

¹ For present purposes I will treat conceptual representations as ones which can enter in syntactic combinations of *indefinite* number and variety. This is not to say that representations in the early visual system do not enter in syntactic combinations with each other, it is just that there is a limit on the number and type of combinations they can enter in.

² Cf. Spelke (1990).

objects go not only with their referent, but also with a representation of location. I argue in section 5 that the way object representations are connected with representations of location--a way implicitly specified in the structure of object-files--can serve as the basis for some of the descriptive concepts above. This makes the connection between early and later representations of objects much more integrated, than on Pylyshyn's account. I examine this issue in section 5, where I present an account of how the object representations in early cognitive processing acquire their content: an account that is closer to Kahneman *et al.*'s view, but which also makes a crucial appeal to the functional roles of object representations. Before that in section 3 I present the evidence for object representations and in section 4 I argue against Pylyshyn.

3. Object Representations In the Early Visual System

Recently, there has been purported evidence that in the early stages of visual processing there are visual demonstratives that tend to refer to salient entities in an observer's environment. For example, in Pylyshyn's Multiple Object Tracking (MOT) experiments³ the test subjects are presented with some number (usually more than five) of what to the subjects looks like identical objects on a computer screen. Following a cue that draws the subjects' attention to four or five of the entities, the subjects are tasked to track them for a short period of time. The entities move randomly but, nevertheless, the subjects are extremely good in tracking four to five of them (87% success rate). The entities are successfully tracked despite changes in their perceivable properties, such as their location, shape, and color, and even despite brief full occlusion. Pylyshyn takes the results to suggest that our visual system employs four to five syntactically differentiated representations--Pylyshyn calls them "FINSTs"--to track, not *properties*, but what typically turn out to be *objects* in the *ambient* environment (as opposed to 'objects' on the retina). Since the FINSTs' contents are singular (the FINSTs are about particular entities in the ambient environment) and context dependent (the FINSTs are about whichever entities happen to be the four or five most salient ones), Pylyshyn takes FINSTs to be a type of demonstratives--visual demonstratives. Further, since only the properties of the tracked entities get encoded in a manner specifying the inter-property relations and since attention is needed to encode such relations, Pylyshyn takes FINSTs to guide attention to their subjects and the subjects' properties.

FINSTs can be treated as the object-files in Treisman and Gelade's (1980) account.⁴ Object-files are representations that organize information registered by feature-maps. Feature-maps, are neural layouts that register a given property and its spatial relation to the observer. There are multiple feature-maps for different properties (texture, location, orientation, etc.). There is also a master-map--a map that registers when feature-maps register information from the *same* region in ambient space. When the master-map detects that multiple feature-maps register information from the same region, this information is stored in object-files--short-term memory banks for salient entities in the ambient environment. Each object-file, then, goes with one representation of a region at a time.

Object-files have two functions. The first one is to refer. Similarly to FINSTs, this function is postulated to explain the following facts: a) objects can be tracked without being identified (e.g. tracking an object in the sky without knowing what it is); b) objects can be

³ Cf. Pylyshyn *et al.* (1988, 2000). See also Scholl and Pylyshyn (1999).

⁴ Cf. Pylyshyn (2007, pp. 37-38): "A convenient way to tell the story of what goes on in MOT is in terms of object files"

tracked despite change in their properties (e.g. tracking a rotating object with an asymmetric shape); and c) objects with the same properties, apart from their location, can be seen as distinct. The second function of object-files is to store and update information about the represented entity. It is postulated to explain effects like the reviewing effect which has been reported by Kahneman *et al.* (1992). The subjects were first given a list of different letters and then a single letter. If both the location and type of the letter matched that of a letter in the list of letters, the subjects' response was faster. This finding was interpreted as showing that the subjects clustered the representation of a letter-type with that of a location and stored them. When the single letter matched the letter-type and location of a letter in the letter list, a new cluster did not need to be encoded, instead the one used for the list was retrieved. The stored information in object-files is updated when the changes in texture, orientation, and location are not too great. If it is, then a new object-file is created. Object-files, prior to being attended, are little more than lists of property representations concatenated with the same representation of location. Attention is needed to extract information about the relations between the various represented properties and may also be needed to track the object-files' referents.

The relation between FINSTs and object-files is that FINST could be treated as the referential part of object-files⁵ (as a result, I will often discuss this question by interchangeably referencing FINSTs and object-files). But there are two views on how FINSTs, or the referential part of object-files works. On the object-file view of Kahneman *et al.* object-files refer by representations of location. Pylyshyn, on the other hand, argues that they refer without the help of location or any other representations.⁶ I will argue below that Pylyshyn's position is untenable.

4. Vs. Pure Causal Theories of Reference

The question of concern in this paper is how representations like FINSTs and object-files get to refer to objects. One common view, defended by Pylyshyn, is that, unless prevented in some manner, any FINST refers to the salient external entity that caused it to be tokened--what normally happens to be an object⁷. In this manner a FINST refers without describing the referent in any manner, even by its location. This is an example of a pure causal account of reference--one where no descriptions are appealed to.⁸ Such accounts easily explain how FINSTs can be visual demonstratives. The referent is a particular--the cause--and what particular it is, is context dependent.

⁵ Cf. Kahneman *et al.* (1992, p. 215-216).

⁶ Cf. Pylyshyn (2001, p. 217): "I claim that no concepts or encodings of properties of the objects are involved in assigning or in maintaining indexes [(FINSTs)]".

⁷ Pylyshyn, however, is often ambiguous with respects to the FINSTs' referents. Sometimes he talks about them as "objects in the distal world" (Pylyshyn (2001, p. 208)) and sometimes as being ambiguous between enduring physical objects and "proximal visual patterns" (Pylyshyn (2001, p. 173)).

⁸ Cf. Pylyshyn (2007, p. 97): "establishing a FINST is entirely a causal process". A problem that has often been raised against such causal accounts is that there are many causes of a FINST--there are many "links" in the causal "chain" of events that lead to a FINST. Pylyshyn (2007, p. 97) notes this and rightly suggests that a plausible solution to this problem is Fodor's (2008, p. 213) triangulation account. According to a charitable interpretation of this account, the link in the causal chain that is the referent of a token symbol is the first intersection of the causal chain with a counterfactual causal chain that leads to the same token symbol, but from slightly different counterfactual perspective.

What is implicit in Pylyshyn's proposal is the assumption that objects are the only causal relata. However, such assumption would be taken as problematic by those who believe that events can also cause or are in general theoretically preferable to objects as the relata in causal relations.⁹ For example, we not only say that a Bosnian assassin (Gavrilo Princip) caused WWI, a case where we have an object (Gavrilo Princip) as a causal relatum, but we also say that the murder perpetrated by Gavrilo Princip caused WWI--a case where we have only events as the causal relata. Further, if it is Gavrilo Princip that is the cause of WWI in 1914, then the assassin should have been a cause of WWI at any point at which he was Gavrilo Princip--that is, at any point of his existence. An easy way to avoid this would be to hold that it is the event of the assassin killing Franz Ferdinand at 1914 that caused WWI, and not the assassin himself.

If both events and objects are causes, then pure causal accounts, like Pylyshyn's, would have FINSTs refer to both objects and events. If it is only events that enter in causal relations, then the referents of FINSTs would be events. However, this is problematic since FINSTs were postulated to explain tracking--identifying an entity as being the same despite changes in its spatio-temporal properties. Events, however, are not entities that can change, as a whole, their spatio-temporal properties. If they could, then, as causes, they would have the same effects in all of their instantiations. For example, consider the FIFA World Cup. People often talk about it as an event that has been played in different continents and so it seems that events can change their locations. However, if we treat the FIFA World Cup as an event, then if we want to say that it caused a large tourist influx to South Africa (in 2010), it should also have caused a large tourist influx to South Africa at any point of time of its existence--including when it was staged in Uruguay (in 1930). An easy way to avoid this result is to treat the FIFA World Cup, not as an event, but as a *type* of events--world soccer events each of which occurs, as a whole, at only one location.¹⁰

The upshot is that if FINSTs are to explain how entities get tracked, they have to be object representations. But then, pure causal accounts cannot be appealed to, by themselves, as an explanation as to how FINSTs refer.

The above arguments, however, depends on two contentious assumptions: that events can be causal relata and that events are distinct from objects. Both have been disputed.¹¹ Still, there is a deeper problem for pure causal theories. If FINSTs get their content purely *via* causal relations, then one would have difficulty objectively distinguishing between representations of objects and representations of universals. This problem arises due to the fact that apart from causal relations with their referents FINSTs also enter in stable co-variations with universals. Pylyshyn takes a FINST to refer to a *particular*--one of the four or five most salient external objects. But FINSTs *also* enter into a stable co-variation with a *universal*--the *property* of being one of the four or five most salient external objects. Now, accounts of how representations of universals refer often appeal at least to stable co-variations between the representations and the referents.¹² But then, the problem is to decide why a FINST be should be treated as a representation of an object, as opposed to a representation of a universal.

⁹ Davidson (1980), Kim (1973), and Lewis (1986b).

¹⁰ Cf. Lewis (1986a, p. 243). See also Dretske (1967).

¹¹ Armstrong (1997) argues that it is facts (e.g. that a is F) that enter in causal relations, as opposed to events. Quine (1985, p. 167): suggests that events can be reduced to objects.

¹² Burge (2010), Dretske (1982), Fodor (1990, 1995).

It is difficult to see how an account that appeals only to the relations between representations and external entities, like pure causal accounts of reference, can justify why one should treat a representation as that of a particular, as opposed to of a universal. For instance, pure causal accounts cannot appeal to syntactic relations to specify the content of a representation. One popular attempt to solve the problem above is to argue that since the mechanism producing FINSTs is *selected for* representing and tracking objects, and not universals, FINSTs are representations of objects.¹³ However, given that it is extremely implausible that a mechanism has been selected for in the *past* to detect a certain *present* object, most likely the mechanism in question operates by detecting a *universal* that tends to be had by objects. But then, we have the above problem again: why is the referent an object, as opposed to a universal instantiated by the object?

To deal with the problems above one needs to go for more than mere causal relations in specifying how a FINST acquires its content. This is what Kahneman *et al.* do. They specify the referent of an object-file *via* an appeal to a representation of locations. As such, their approach has a descriptive element--the location representation that helps specify the object-file referent. Below I argue that while their approach is on the right track, it needs to be modified to also include an appeal to the functional roles of representations of objects.

5. Predications stored in memory

The problems for pure causal accounts are a reason to appeal not only to causal roles, but also to descriptions in determining the referent of FINSTs and object-files.¹⁴ Remember that FINSTs were postulated as demonstrative representations and the initial appeal of pure causal accounts was that they could easily account for the demonstrative element of a FINST's nature. But descriptive accounts of reference can explain just as easily as pure causal ones how FINSTs can be demonstratives, as long as they are not pure. Roughly, on such accounts the referent of a FINST is the entity that a) satisfies a description concatenated with the FINST; and b) is in the cause of a FINST. Since what the description and the FINST's cause are is context dependent (in different contexts different entities would give rise to FINSTs and property representations), the FINST's referent is context dependent--as a referent of a demonstrative should be.

The referent of a demonstrative should also be same one in every counterfactual circumstance of evaluation. For example, in evaluating the utterance, "that man is the US president", uttered while I am pointing to Obama, the referent of "that man" could only be Obama, whether or not I evaluate it with respect to a world where Obama is the US president. In other words, the referents of demonstratives are rigidly specified. To accommodate this, a descriptive account of how demonstratives refer cannot be a purely descriptive one or else it would suffer from well-known problems.¹⁵ For instance, the referent of "the current president of the US" could have been Romney, when evaluated with respect to a possible world where

¹³ Neander (1995), Burge (2010, p. 322-326).

¹⁴ Pylyshyn has argued that descriptions are not required to represent and track objects (*cf.* Pylyshyn and Storm's (1988)): using a computer simulation of tracking, where focal attention moved serially among the tracked objects and updated a record of their locations, the success rate was only 30%, compared to that of humans--about 90%. But, as Hill (2008) points out, Pylyshyn's empirical argument does not work against accounts where tracking is done in virtue of location representations that get updated in parallel (perhaps by focal attention glomming onto several FINSTs at once). This suggests that descriptions can be used in tracking and representing objects.

¹⁵ *Cf.* Kripke (1980, p. 3-15, 48-78).

Romney won the election. In other words, while demonstratives are rigid designators, descriptions are not. Thus, if one is to give an account of demonstratives *via* descriptions, one would need to rigidify the descriptions. A simple way to rigidify them would be to include a causal element. Roughly, the causal element specifies that the referent is contained in the actual cause of a description. In this manner a rigidified description would have, just like a demonstrative, the same referent in every counterfactual circumstance of evaluation.¹⁶

So, given that descriptive accounts seem more likely than purely causal ones to offer a plausible account of the demonstrative nature of FINSTs and of how FINSTs and object-files refer to objects, we reach the conclusion that FINSTs, contrary to what Pylyshyn claims, do refer *via* descriptions. That is, we can conclude that FINSTs refer *via* representations of universals. But then, if FINSTs are related to representations of universals and refer *via* them, they start to look much like Kahneman *et al.*'s object-files and not merely as the referential part of object-files and one that is independent of any property representations stored in the "files". As a result, and this is what I will mean to do from now on, we can now treat FINSTs and object-files as the same type of representation--one that I will call a "visual index".

We now can turn to the question of what the general form of such a descriptive account would be like. In what follows I will argue that to have representations of objects the descriptive element needs to include stored in memory predications containing location representations. In section 5.1 I make the case for the appeal to predications and then suggest that there is good reason to believe that we have predications at the early stages of cognitive processing. In section 5.2, I point out that an appeal to predications suffices to distinguish representations of particulars from those of universals. Nevertheless, as I point out in section 5.3, such predications are not enough to specify the representations of particulars as representations of objects, as opposed to of other particulars. Finally, in section 5.4, I argue that we also need to appeal to memory to explain why visual indexes are object representations.

5.1. Predication

The upshot from the discussion above is that to account for how visual indexes refer we have to appeal to descriptions. But how do we appeal to descriptions? How can we get a representation of a universal to attribute its content to a representation of a particular? We cannot simply concatenate representations--this would not represent any relation between them, let alone represent attribution. We need *predication*. But can we talk about predication when it comes to processes in the early visual system? I will argue now that it is not without precedent and not implausible.

For example, Clark (2000, 2004) has argued that we need to talk about predication in the early visual system in order to solve the binding problem: the problem of determining which representations of properties go with which representation of particular (so that we can distinguish a red square next to a green circle from a red circle next to a green square). Clark's proposal is that we need representations of two syntactically different types: subjects and predicates, where the subjects "anchor" the predicates. Clark, *via* an appeal to Strawson (1959,

¹⁶ Another reason for keeping a causal element in descriptive accounts is that if there is no causal connection between the external world and mental states, then the mental states, by themselves, are not sufficient to determine anything *about* the external world--the same mental states can exist in radically different external worlds. That much is true at least for the mental states in the early stages of sensory processing. *Cf.* Putnam (1975).

1974), points out that the relations that obtain between subjects and predicates--relations analysed by Strawson--also obtain between representations of location and the property representations found in feature maps. In much the same way, as I will now argue, we can treat visual indexes and the representations of properties stored in them as forming a predication because they enter in many of the relations Strawson specifies between arguments and predicates.

For the purposes of this paper I will assume that something like Strawson account¹⁷ of how subjects and predicates differ from each other is on the right track and will just present one of the ways that he describes, because it is going to be key in explaining what distinguishes a representation of an object from representations of other particulars. Strawson points out that there is the following asymmetry with regards to subjects and predicates: with regards to predicates there are incompatibility groups, whereas there are no such groups with regards to subjects. An incompatibility group is a range of predicates such that if a predicate out of that range is concatenated with a subject, then, regardless of what the subject is, no other predicate out of this range can be concatenated with the subject (e.g. 'square', 'triangular', and other shape predicates form such incompatibility groups--if a subject is concatenated with 'square' it cannot be concatenated with 'triangular'). There are no incompatibility groups with regards to subjects: it is not the case that there is a range of subjects such that if a subject out of that range is concatenated with a predicate, then regardless of what the predicate is, no other subject out of this range can be concatenated with the predicate (e.g. one subject being concatenated with 'square' does not preclude another from also being so concatenated). Strawson's point, then, is that what it means to be a predication just is for representations to enter in the asymmetric relations that he has specified. Intuitively, if representation 'P' restricts how representation 'e' behaves, but not *vice versa*, then 'P' and 'e' act like a universal and an instantiation of it--the universal determines how its instantiation causally interacts with the world, but not *vice versa*.

Assuming that Strawson is right about predicates, but not subjects, entering into incompatibility groups, what evidence is there that visual indexes and property representations enter in such relations? According to Kahneman *et al.*'s view of object-files, when properties are detected to be from the *same location*, they and their location get encoded in one object-file. Each object-file, then, goes with only one representation of a location. What we have is that representations of locations enter into incompatibility relations--once one is attached to an object-file, another representation of location, at that time, cannot be also so attached to that object-file. While location representations enter into incompatibility groups, object-files do not. The fact that one object-file contains the representation 'red', say, does not entail that another one cannot also do so.¹⁸

¹⁷ Cf. Strawson (1959, p. 167-170) and (1974, p. 103).

¹⁸ The other asymmetries that Strawson points out can also be seen to obtain with respect to object-files and the property representations stored in them. First, we have what, for brevity's sake, I will call "instance relations": while different predicate instances can be concatenated with one subject (e.g. 'Socrates' smile' and 'Socrates' frown' can both be concatenated with 'Socrates'), only one subject can be the subject of a predicate instance (e.g. 'Socrates' smile' can be concatenated only with 'Socrates'). We can see evidence for such relations in object-files and feature-maps. While different representations of properties at locations (representations of properties from different feature-maps) can go with one object-file, only one object-file can contain any particular representation of a property at a location (at a time). The reason is that a representation of a property at a location is also a representation of a location and each such representation is concatenated with only one object-file. Second, with regards to predicates there are involvement relations, whereas there are no such relations with regards to subjects. Roughly, an involvement relation is a relation between predicates such that one implies the other, but not *vice versa* (e.g. the determinate/determinable relation: being scarlet implies being red, but not *vice versa*). There are no such involvement relations with regards to subjects: a subject being scarlet does not entail that another subject is

Given the evidence for the asymmetric relations that Strawson specifies between representations in the early visual system, we can justify the appeal to predication at those early stages of cognitive processing. This removes a major hurdle for descriptive accounts of how visual indexes refer, since it allows their content to be specified by the predicates they are concatenated with. However, given that universals can have other universals as attributes (as in second-order logic), more work is needed to have visual indexes refer to objects, as opposed to universals. I turn to dealing with this problem below.

5.2. Particulars or Universals?

I am going to argue now that if we can appeal to predication, we can determine why visual indexes represent particulars, as opposed to a universal. On the face of it, just because a visual index enters in a predication with a representation of a universal does not make the index a representation of a particular--property representations, after all, can enter in predications with other property representations, as well. However, what distinguishes property representations from representations of particulars is that the former can play the role of an argument and a predicate, while the latter can only play the role of an argument. Intuitively, just like a particular is not an entity that is an attribute to anything, a representation of a particular cannot attribute its content to the referent of another representation. For example, it makes no sense to say, "Superman has (the property) Clark Kent" because "Clark Kent" is not a property (unlike the property of *being* Clark Kent). The upshot, then, is that what it means for a representation to be one of a particular, as opposed to a universal, is for it to be only capable of playing the role of an argument in predications. Essentially, this is what grounds all of Strawson's asymmetric relations between predicates and arguments. If there were symmetry, then both arguments and predicates would be able to play each other's roles and there will be no syntactic difference between the two.

Is there evidence that visual indexes can only play the role of arguments in predications? As pointed out above, Kahneman *et al.*'s data about object-files shows that there are asymmetries, of the sort that Strawson specifies, between the roles of the representations stored in object-files and the object-files themselves. This means that object-files cannot play the roles of predicates and are therefore representations of particulars.

Knowing that visual indexes represent particulars, however, is only half the journey: we still need to explain why a visual index represents a particular type of a particular--why it is an object, say, as opposed to a region, trope, or an event. This is the topic of the next two sections.

5.3. Objects, Regions, Events, or Tropes?

I have argued above that visual indexes are arguments and represent particulars in virtue of entering in asymmetric relations, of the sort that Strawson specifies, with

also scarlet. We can again see evidence for such relations in object-files and feature-maps. There are different feature-maps for different determinables--colors, shapes, motion, etc.. Within each feature-map we have neurons that respond to a particular determinate of the corresponding determinable. For example, within the "red" feature-map we have neurons that respond to being scarlet. This means that an active neuron from the "red" feature-map (e.g. the one for scarlet) would entail that something red is seen, but just seeing red would not entail that the "scarlet" neuron is active. There don't seem to be corresponding involvement relations with respect to object-files: the fact that one object-file contains a representation of red does not entail that another one must also have it.

representations of universals. What else is needed to have visual indexes refer to objects? I am going to argue now that simply attributing a property to the referent of a visual index would not suffice to make the referent of a visual index an object. *Prima facie*, we should be able to specify an object as a visual index's referent by appealing to predicates like 'object' or, perhaps more realistically, given that such predicates are unlikely to exist at the early stages of visual processing, to representations of properties like being bounded, unified, and persisting through time? These are, in fact, the criteria that Spelke (1990) has demonstrated children use to detect objects. This may seem to work, given that the property of persisting through time distinguishes objects from events, regions, and tropes. While regions and events can both be described as unified and bounded, objects, but not regions and events can be wholly instantiated at different spatio-temporal locations. So, perhaps, if a visual index's referent is described as being bounded, unified, and persisting through space-time, then the referent would not be a universal, since universals do not persist through space-time and so cannot be tracked. It wouldn't be an event either, since events cannot be wholly instantiated at different spatio-temporal locations and so cannot be said to persist through space (I will discuss an obvious rebuttal below).¹⁹ It would seem that the referent cannot be a region, because the tracked entity changes its *spatio*-temporal coordinates and regions cannot change those. Finally, it would seem that the referent cannot be a property instance (a trope), since a property instance is an object's having a property *at a time*²⁰, and the tracked entity changes its *spatio-temporal* coordinates. Then, the referent, being a particular, would seem to have to be an object.

However, any approach that appeals to simple attribution of a property would fail because such attribution is ambiguous between "property at" and "property of". Suppose that a visual index is concatenated with predicates representing the properties of being bounded, unified, and persisting through space-time. Still, it remains unclear whether the referent is an object that *has* all these properties or whether it is a region *at* which the properties are instantiated.

Another problem is that of determining in what sense a visual index's referent *persists*. There are, roughly, two views on what persistence means. According to the 3D view entities do not have temporal parts and persist just in case they are wholly located at different temporal moments. On a 4D view, on the other hand, entities have temporal parts and persist in the sense of having different temporal parts at different temporal moments. Now, on a 3D view one can easily distinguish between objects and events: the former persist in the sense of being wholly located at different spatio-temporal locations, while the latter do not. However, on a 4D view both objects and events both have temporal parts and persist in the sense of being spread out in space and time. Thus, on a 4D view, just labeling the referent of a visual index as persisting would not automatically single out an object, as it would on a 3D account.²¹

Finally, with respect to visual indexes, it is implausible that we can appeal to representations of properties like the property of being a persisting entity. Visual indexes contain representations of colors, locations, and textures--much "simpler" properties than the property of being a persisting entity. Thus, one can question whether, even if attributing persistence would serve to single out an object as the referent of a visual index, such an approach is viable at the early stages of cognitive processing.

¹⁹ Cf. Dretske (1967).

²⁰ Cf. Funkhouser (2006, p. 14).

²¹ Cf. Quine (1985, p. 167). As Quine (1985) has noted, one may have difficulty differentiating between objects and events in a 4D view: both can be thought of as "the material content of any portion of space-time".

The upshot is that even if we describe the referent of a visual index as one that persists through time, still, we succeed only in weeding out universals as potential referents (universals do not persist in either the 3D or 4D sense), but not regions and events. What to do? In what follows I propose that *via* an appeal to memory and implicit representation of identity, we can isolate objects as the referents of visual indexes.

5.4. Weeding Out Regions, Events, and Tropes

My strategy in explaining why visual indexes refer to objects, as opposed to other particulars, is to appeal to the visual indexes' functional roles. I propose that to specify why a subject refers to an object we need to appeal to a) incompatibility groups of spatio-temporal predicates; and b) predications stored in memory. The appeal to incompatibility groups of *spatio-temporal* predicates is to give an account of how entities are represented as bounded²². Once the referent of a subject is represented as bounded, then one can appeal to the fact that a region, trope, or an event cannot be bounded at two different spatio-temporal regions, to explain why a subject refers to an object and not to one of those other types of particulars. I explain these two elements of the proposal below.

Suppose that a subject is concatenated with a spatio-temporal predicate in a way that entails that the subject cannot be concatenated with any other spatio-temporal predicate. This would mean, by the definition of incompatibility groups above, that the predicate enters in an incompatibility group with other spatio-temporal predicates. This fact--of a spatio-temporal predicate entering into an incompatibility group--means also that the referent of the subject is represented as *bounded* by the region specified by the spatio-temporal predicate. To see this, suppose that 'b' is an argument concatenated with the predicate 'R', where this predicate enters into incompatibility groups and represents the property of being a circular region in ambient space with a certain radius and a center specified in observer-relative coordinates. That the predicate enters into incompatibility relations means that the referent of 'b' is represented as *not being related to any other region at t*. If b is represented as occupying space-time *outside* R, then there would be another space-time region that b is represented as being at--one outside R (at least partially). If b is represented as being *within* R, but not filling it, then there would also be another space-time location that b is represented as being at--one within R. This means that when 'R' enters into incompatibility relations, b is represented as being: a) *within* the region R; and b) *filling* the region R. In other words: if it is represented as *bounded* by this region.

Having a subject related with a spatio-temporal predicate that enters into incompatibility groups is obviously not sufficient to specify the subject's referent as an object, since regions can also be bounded. However, the situation changes if we can have a particular being represented as bounded at two different space-time regions. To represent a particular in this manner means to represent it as persisting through space-time where at any moment the particular is bounded (i.e. wholly located) at that moment and therefore does not have temporal parts in other moments. But that means that a system that represents particulars in this manner represents them as persisting in accordance with a 3D account of persistence, where, roughly, a particular persists from t_1 to t_2 just in case it is wholly located at both moments. This allows the

²² Cf. Spelke (1990, p. 1): "Infants divide perceptual arrays into units that move as connected wholes [i.e. as bounded], that move separately from one another, that tend to maintain their size and shape over motion, and that tend to act upon each other only on contact. These findings suggest that a general representation of object unity and boundaries is interposed between representations of surfaces and representations of objects of familiar kinds."

system to discriminate between objects and regions, events, or property instances. Only objects can be wholly instantiated at two different space-time regions. Regions, events, and property instances are all defined in terms of their spatio-temporal locations. That is, they can never abandon their spatio-temporal locations. Objects, on the other hand, not being defined in terms of their spatio-temporal locations, can wander in space-time.²³

How would an entity be represented as bounded at two different space-time regions? The goal is to have a subject 'b' be related to two different spatio-temporal predicates, where the predicates enter into incompatibility groups. However, because of the incompatibility relations we cannot simply have 'b' be related with two spatio-temporal predicates--after all, if a spatio-temporal predicate enters into incompatibility relations with other spatio-temporal predicates, then its subject cannot be related to any other spatio-temporal predicates, but it. This seems to prevent us from relating 'b' with two different spatio-temporal predicates at the same time. But this problem can be avoided if we a) assume that spatio-temporal predicates go with a temporal index (e.g. 'R_t'); and b) have an incompatibility group containing 'R_t' be defined as one such that if a subject is concatenated with 'R_t', no other spatio-temporal predicate with the same temporal index (i.e. t) can be concatenated with the subject. For example, 'b' cannot be related at t₁ to any other t₁-indexed space-time predicate, but 'R_{t₁}'. Nevertheless, 'b' can still be related at t₁ to spatio-temporal predicates with a temporal index different from 't₁', for example the predicate 'R_{t₂}'. This allows for 'b' to be related to two different spatio-temporal predicates where the predicates enter into incompatibility groups.

The upshot is that we have predications like 'R_tb', where 'R_t' is a spatio-temporal predicate representing a region at a time. We also have conjunctions of predications like 'R_{t₁}b & R_{t₂}b', where R_{t₁} and R_{t₂} both enter into incompatibility groups. That a spatio-temporal predicate goes with a temporal index means that the predications formed with such predicates have to be stored in *memory*. This means that for representations of objects we need not only predications with spatio-temporal predicates where the predicates enter into incompatibility groups, but also that these predications are stored in memory.

Because the predications contain spatio-temporal predicates that enter into incompatibility groups, the referent of 'b' is represented as being bounded by at least two different spatio-temporal regions. This, as explained above, means that it represents an object, as opposed to a region, event, or a property instance. To exclude the possibility of 'b' representing two objects as being bounded by the same region, we may also need to add that the predications are definite descriptions. Then, each predication would represent only one object as satisfying the predication.

Is there evidence for predications stored in memory when it comes to the early visual system? I have already argued that there is evidence for visual indexes entering in predications. Evidence for predications stored in memory is ample. Some examples come from Pylyshyn's

²³ The above argument crucially assumes that a *single* particular gets represented as being bounded at two different space-time regions. But why assume that a representation at t₂ refers to an entity as the *same* entity as one represented at t₁? In other words, why assume that we have a representation of identity here? One reason is that we have the same token symbol used--the token symbol that is used at t₁ is used at t₂, as well. But why couldn't the fact that the same symbol is used represent the fact that two distinct entities at t₁ and t₂ are stages or temporal parts--ones that are bounded at their location--of the same entity? Well, temporal parthood is a triadic relation: x is a part of z at t. It involves two representations of particulars. In the case of FINSTs and object-files, however, we have just one representation of a particular (the FINST/object-file) concatenated with predicates. This means that we simply don't have the sufficient building blocks for a representation of a parthood relation and so we can assume that no stages or temporal parts are represented.

tracking experiments and cases of a reviewing effect. For example, in his tracking experiments that involve tracking objects through occlusion, Pylyshyn observed that tracking success depended on how close the location of an object's disappearance behind an occluder was to the location of re-appearance. Pylyshyn concluded that representations of locations are created when objects disappear and retrieved after the occlusion in order to associate a prior FINST with the after-occlusion object.²⁴ Kahneman *et al.*'s evidence of a "reviewing effect", presented in section 3, offers further support that object-files are stored in memory.

In conclusion, given that visual indexes enter in predications stored in memory and are concatenated with spatio-temporal predicates that enter into incompatibility groups, we can treat them as representations of objects.

6. Conclusion

The question examined in this article was how representations of objects at the early stages of visual processing acquire their content. I have argued that two conditions need to be satisfied: a) the representations must enter in stored in memory predications with spatio-temporal predicates; and b) the spatio-temporal predicates must enter into incompatibility groups. This account has the benefit of not depending on a representation of the property of being a persisting entity to describe an entity as persisting and as an object. Instead it shows how an entity can be implicitly represented²⁵ as persisting--in virtue of how a representation of a particular is related with spatio-temporal predicates stored in memory.

²⁴ Cf. Pylyshyn (2007, p. 80): "...our assumption is that the disappearance itself causes locations to be conceptualized and stored in memory." See also Pylyshyn (2007, p. 40): "[it] seems at least that when tracked targets disappear there is a record of where they were when they disappeared."

²⁵ Information is implicitly represented when an inference is required to extract it.

Bibliography

- Armstrong, D. (1997). *A World of States of Affairs*. New York, NY: Cambridge University Press, 1997.
- Burge, T. (2010). *Origins of Objectivity*. New York, NY: Oxford University Press Inc., 2010.
- Clark, A. (2004). Feature-Placing and Proto-Objects. *Philosophical Psychology*, 17(4), December 2004, 451-477.
- Clark, A. (2000). *A Theory of Sentience*. Oxford University Press.
- Davidson, D. (1980). Actions, Reasons, and Causes, originally in 1963, *Essays on Actions and Events*. Oxford: Clarendon Press, pp. 3-19.
- Dretske, F. (1967). Can Events Move? *Mind*, New Series, Volume 76, Issue 304 (Oct., 1967), pp. 479-492.
- Dretske, F. (1982). *Knowledge and the Flow of Information*. Cambridge MA: MIT Press.
- Fodor, J. (2008). *LOT 2: The Language of Thought Revisited*. Oxford, England: Oxford University Press, 2008.
- Fodor, J. (1995). *The Elm and the Expert*. Bradford Books/MIT Press, 1995.
- Fodor, J. (1990). *A Theory of Content and Other Essays*. Cambridge, MA: MIT/Bradford, 1990.
- Funkhouser, E. (2006). The Determinable-Determinate Relation. *Nous*, 2006, Vol. 40; # 3, pp. 548-569.
- Hill, C. (2008). Review of *Things and Places*. *Philosophy Reviews*. Notre Dame Philosophical Reviews.
- Kahneman, D., Treisman, A., and Gibbs, B. (1992). The Reviewing of Object Files: Object Specific Integration of Information. *Cognitive Psychology* 24, 175-219 (1992).
- Kim, J. (1973). Causation, Nomic Subsumption, and the Concept of Event. *Journal of Philosophy* 70, pp. 217-36.
- Kripke, S. (1980). *Naming and Necessity*. Harvard University Press, 1980.
- Lewis, D. (1986a). Causation. *Philosophical Papers 2*. Oxford: Oxford University Press, pp. 159-213.
- Lewis, D. (1986b). Events. *Philosophical Papers 2*. Oxford: Oxford University Press, pp. 241-69.
- Neander, K. (1995). Misrepresenting and Malfunctioning. *Philosophical Studies*, 79: 109-41.
- Pylyshyn, Z. (2007). *Things and Places: How the Mind Connects with the World*. Bradford Books/MIT Press, 2007.
- Pylyshyn, Z. (2001). *Seeing and Visualizing: It's Not What You Think*. Bradford Books/MIT Press, 2001.
- Pylyshyn, Z. and Sears, C. (2000). Multiple Object Tracking and Attentional Processing. *Canadian Journal of Experimental Psychology*, 2000, 54 (1), pp. 1-14.

- Pylyshyn, Z. and Storm, R. (1988). Tracking Multiple Independent Targets: Evidence for a Parallel Tracking Mechanism. *Spatial Vision*, 3(3), pp. 1-19.
- Quine, W. (1985). Events and reification. *Actions and Events: Perspectives on the Philosophy of Davidson*. Blackwell (1985), pp. 162-171
- Scholl, B., and Pylyshyn, Z. (1999). Tracking multiple items through occlusion: Clues to visual objecthood. *Cognitive Psychology*, 38(2), pp. 259-290.
- Spelke, E. (1990). Principles of Object Perception. *Cognitive Science*, 14, 29-56 (1990).
- Strawson, P. (1959). *Individuals: An Essay in Descriptive Metaphysics*. Routledge, Taylor and Francis Group, 1959.
- Strawson, P. (1974). *Subject and Predicate in Logic and Grammar*. Ashgate Publishing Limited, 1974.
- Treisman, A. and Gelade, G. (1980). A feature integration theory of attention. *Cognitive Psychology*, 12, pp. 97–136.