

Forthcoming in Review of Philosophy and Psychology

<http://link.springer.com/article/10.1007/s13164-015-0239-2>

The final publication is available at link.springer.com

Retinal images and object files: towards empirically evaluating philosophical accounts of visual perspective*

Assaf Weksler

The Open University of Israel

Ben Gurion University of the Negev

Assaf.weksler@openu.ac.il

Abstract: According to an influential philosophical view I call “the relational properties view” (RPV), “perspectival” properties, such as the elliptical appearance of a tilted coin, are relational properties of external objects. Philosophers have assessed this view on the basis of phenomenological, epistemological or other purely philosophical considerations. My aim in this paper is to examine whether it is possible to evaluate RPV empirically. In the first, negative part of the paper I consider and

* I am grateful to audiences at the Consciousness and Intentionality conference at the University of Salzburg (2013), the annual conference of the New Israeli Philosophical Association at Bar Ilan University (2013), and the Rutgers-Jagiellonian Conference in Cognitive Science at the Jagiellonian University of Krakow (2014). I would like to thank Sara Aronowitz, Hagit Benbaji, Zoe Gutzeit, Amir Horowitz, Arnon Keren, Arnon Levi, Olivier Massin, Takuya Niikawa, Thomas Papathomas, Eva Schmidt, Itay Shani, Dan Shargel, Błażej Skrzypulec, Ronen Tsdaka, Benjamin Young and an anonymous referee for *The Review of Philosophy and Psychology* for helpful feedback on earlier versions of this paper. Special thanks go to Susanna Schellenberg and Shimon Ullman, whom I treated as paradigmatic representatives of the two main issues the present work addresses, namely RPV and vision science, respectively, and whose generous feedback and encouragement have convinced me that the present project is worth pursuing. I am grateful to Yaron Senderowicz, who supervised my PhD thesis, in which the present research is rooted, for teaching me the difference between a good philosophical idea and good philosophical work, and to Amir Horowitz, my supervisor during my post-doctoral stay at the Open University of Israel, during which the present research was carried out, for all the help and encouragement. This work has been supported by a post-doctoral fellowship plus travel stipends from the research authority of the Open University of Israel (on the basis of Amir Horowitz’s ISF grant no. 419/11).

reject a certain tempting way of doing so. In the second, positive part of the paper I suggest a novel way of evaluating RPV empirically, relying on the influential object files framework.

1. The relational properties view (RPV)

The objects around us usually, in normal conditions, look to have (at least roughly) the shape they actually have: a tilted coin, for example, looks circular. Call properties of this sort “objective shape”, and more generally “objective properties” (covering objective size, objective angle between edges, etc.). Now there is a sense in which ordinary objects present certain shapes that are dependent on the location of the subject and its line of sight. For example, a titled coin looks elliptical from my perspective. Call properties of this sort “perspectival shapes” and more generally “perspectival properties” (covering perspectival size, perspectival angle between edges, etc).

According to a leading philosophical view, perspectival properties are relational properties¹ of external objects (Cohen 2010; Harman 1990; Hill 2009; Lycan forthcoming; Noë 2004; Schellenberg 2008; Tye 2002; Brogaard 2010 accepts this view except she holds that the properties are *centered*, rather than relational; Fish 2009 accepts this view but construes it in terms of *acquaintance* rather than in terms of representation.). Call this “the relational properties view” (RPV). From this point on I mainly focus on perspectival shape, but my discussion equally applies to all other perspectival properties.

According to RPV, then, perspectival shapes are relational properties of external objects. An example of such a property I will use throughout is the property of projecting an elliptical shape on an interposed plane perpendicular to the line of sight (see Noë 2004, pp. 81-82; Tye 2002, p. 79), or in short, the property of having an elliptical projection. Cohen (2010) and Hill (2009) offer other candidate properties; the issues I raise here equally apply to them.

¹ Relational properties are properties of being in a relation with something. Examples include the properties of being an aunt, of causing a fire, of being inside a car, and of owning a pet.

What are the philosophical reasons for thinking that perspectival shapes are relational properties? Let me briefly present the most familiar ones. If we do away with relational properties, appealing instead to ordinary shape properties, then since the tilted coin is not literally elliptical, we need to say either (a) that the experience (of the tilted coin) veridically represents *something else* as literally elliptical, namely an internal object (e.g., a retinal region), or (b) that the experience *misrepresents* the coin as literally elliptical. Both options are philosophically unattractive. Let me consider them in turn. The first option amounts to resurrecting the (or at least something closely related to the) sense-datum theory, which has long ago been discarded by mainstream philosophy of perception on metaphysical and epistemological grounds, and which moreover appears to conflict with the phenomenology, especially with the transparency of experience. As Hill (2009) puts it,

“when we consider our visual experience introspectively, we find no grounds for saying that awareness of ordinary objects is mediated by awareness of objects of some other kind. It is not mediated by awareness of internal mental objects, and it is not mediated by “extra-ordinary” physical objects, such as retinal images or packets of structured light in the area immediately before one’s eyes. Rather, it seems that we open our eyes and ordinary objects are simply there. Now if appearances [i.e., perspectival properties] are properties that we are aware of in visual experience, and the only objects of visual awareness are ordinary physical objects, then it must be true that appearances are properties of ordinary objects” (2009, pp. 143-144).

As I understand it, Hill’s idea is that phenomenological reflection tells us that the only objects we see are external objects. Assuming phenomenological reflection is reliable, we can conclude that we do not see internal items (such as retinal regions), contrary to the teachings of the sense-datum theory. Adding the plausible assumption that seeing is always a case of an object appearing to have some property, it

follows that all the properties we see – perspectival properties included – are properties of external objects.

According to the second option (call it The Error Theory), the experience represents the titled coin as both literally elliptical and circular (Lycan, 1996, ch. 7). However, because the titled coin is not literally elliptical, this option entails the existence of massive perceptual error, that is, it entails that almost every visual experience represents its object as having a property it in fact lacks. Philosophers tend to find this consequence hard to accept (cf. Overgaard 2010, sections 4, 5; Schellenberg 2008, pp. 65-67; Lycan, forthcoming, now says he has abandoned his original view in favor of Schellenberg’s version of RPV). To simplify the discussion ahead, I shall assume, with the aforementioned philosophers, that the Error Theory is false; that is, I shall assume that experiences of perspectival properties are usually veridical, in ordinary circumstances.²

Philosophers typically treat the question as to which of the various views of perspectival properties is correct as a philosophical question at its core, which does not incorporate an empirical question to a significant extent. More precisely, as far as I can tell, all the leading publications that either defend or

² One might suspect that RPV excludes the possibility of perceptual error and that such a consequence appears to be as unacceptable as massive perceptual error. A reason for thinking that RPV excludes the possibility of perceptual error is that in order to misperceive the relational property of projecting a certain shape (or size), there should be a difference between the shape (size) the object actually projects and the shape (size) the object appears to project, and it is difficult to think of a case of this sort. Response: Palmer (1999, p. 316) describes an illusion (“the hallway illusion”) in which a certain cylinder’s (call it “A”) *perspectival* size appears to be approximately half of the *perspectival* size of a different cylinder (call it “B”), yet the size of the retinal image cylinder A produces is about a third of the size of the retinal image cylinder B produces. If *perspectival* size is the relational property of projecting a certain size, it follows that the present case involves a misperception of A’s *perspectival* size. Thus, RPV does not exclude the possibility of perceptual error.

attack RPV (except for some remarks by Burge, discussed below) make their case on epistemological, phenomenological or other purely philosophical grounds. Sometimes philosophers working on this issue mention scientific results, but the role these results play in their argument is minor. For example, both Cohen (2010) and Hill (2009) mention vision-scientific findings, but these do not serve to support RPV over alternative views. Rather, they only serve the role of filling in the details of the antecedently established RPV. For example, the scientific findings constrain the specific relational property Cohen and Hill propose.

Is it possible to empirically evaluate RPV? In section 2 I explore and reject a certain tempting way of doing so. In section 3 I develop a new empirical approach to evaluating RPV, relying on the object files framework.

2. Does vision science straightforwardly conflict with RPV?

Many vision scientists study the question of how representations of objective properties are calculated on the basis of representations of perspectival properties. These scientist tend to speak of these perspectival properties using phrases such as “retinal images”, “retinal projections”, and more specifically: “retinal size”, “retinal shape”, “retinal angle”, “retinal velocity” etc., suggesting that these scientists think of perspectival properties as intrinsic properties of regions in the retina (or of patterns of light striking the retina, I ignore this difference henceforth), rather than as relational properties of external objects (see, for example, Biederman 1987; Feldman & Tremoulet 2006; Feldman 2007; Marr 1982; Palmer 1999; Pylyshyn 2007; Todorovic 2002; Treisman 1999; Ullman 1979, 1996). Call this view RET. RET is a version of the sense-datum theory, and as such conflicts with RPV, apparently making portions of the vision-scientific literature incompatible with RPV. So this seems to be a straightforward empirical claim that refutes RPV and supports the sense-datum theory.

In response, a proponent of RPV might say this³: we should not interpret scientists' talk of "retinal shape" as aimed at accounting for the phenomenology of perspectival shape. In other words, we should not attribute to vision scientists acceptance of RET. We should instead interpret them as speaking of unconscious representations of perspectival shape, claiming that *these* are of retinal shapes (call this claim URET, for "Unconscious RET"), thereby leaving the question about how to account for conscious experiences of perspectival shape open. The way some vision scientists speak about retinal images often invites this interpretation. For example, consider the widely studied phenomenon of perceptual constancy. Palmer (1999, p. 312) defines it (on some occasions) as "the fact that people veridically perceive the constant, unchanging properties of external objects *rather than* the more transient properties of their retinal images" (Palmer 1999, p. 312, my emphasis). Thus, in cases of perceptual constancy so defined, people do not see properties of retinal images. In other words, in the cases in question, visual representations of properties of retinal images are *unconscious*. Thus, scientists who focus on perceptual constancy (as Palmer defines it) do not try to account for the conscious perception of perspectival properties at all. Hence, they are not committed to RET but rather only to URET.

URET is compatible with the following variant of RPV: on the basis of unconscious representations of retinal shapes (and other retinal features), representations of relational properties of external objects are calculated, and these are conscious, hence they account for the (conscious) perspectival shapes objects look to have. On this picture, there are two kinds of perspectival properties represented in visual processing of a tilted coin: there are relational properties such as "having an elliptical projection" and intrinsic properties of retinal regions such as (literal) "ellipticality". Only the former are conscious. Call this the "dual perspectival view", or DP.

Arguably, something like DP is what Burge targets in the following remark:

³ Susanna Schellenberg (personal communication) has suggested to me a response of this sort (as a first response).

“I believe that it is a philosophical and scientific mistake to regard any objective ‘perspectival’ properties [i.e., relational properties as suggested by RPV], such as perspectival size, shape, color, as among the objective environmental entities seen, unaided by background theory. [...] Vision science does not take perspectival appearances as perceptual referents. *I see no need for it to do so*” (2005, fn. 19, my emphasis, see also Burge 2010, pp. 391-392 for similar remarks).

Burge’s remarks are sketchy, and he does not develop them into a full-fledged argument. In any event, his remarks suggest to me the following argument against DP. DP implies that, when one looks at a tilted coin then, in addition to a representation of an elliptical retinal region, the visual system represents the coin as having an elliptical projection. The second representation, on this view, is conscious, whereas the first is unconscious. The problem is that whatever calculation the visual system needs to make (such as calculating the objective shape of the coin), the first “elliptical” representation is just as good as the second, making the second redundant. The result is that the conscious experience of the coin as looking elliptical is redundant, which is *prima facie* implausible. Moreover, some proponents of RPV (e.g., Noë 2004; Schellenberg 2008) explicitly argue that conscious experiences of perspectival properties play a role in allowing us to see objective features of objects. If they are correct, the claim that conscious experiences of perspectival properties are redundant is false. Hence DP is significantly unattractive, especially for proponents of RPV, so I think we can safely ignore DP, together with its emphasis on URET.

Let us return to RET. I grant that RET refutes RPV. So it appears that we have located an empirical claim that refutes RPV. However, as I argue next, this is not right, because RET is not a serious scientific claim, hence it is not legitimate to rely on it to test (and refute) RPV. By saying that RET is not a serious scientific claim (henceforth NSS) I mean that scientists adopt RET – when they do – not on the basis of argument, empirically based or otherwise, but rather automatically and unreflectively.

In support of NSS I present two considerations. The first is that vision scientists do not publish papers attacking or defending RET. We simply find scientists adopting RET (and sometimes, more rarely, adopting RPV) in incidental remarks without any argument or justification, usually in the introductory parts of their work.

To this one might reply that vision scientists adopt RET without argument because RET simply follows straightforwardly from a basic claim shared by all vision scientists, namely that the visual processes computing representations of objective properties begins with a 2D registration of light intensities by cells in the retina. These cells register the retinal image, i.e., the distribution of light intensities over the retina. When looking at a tilted coin (in normal viewing conditions), there is an elliptical pattern in the retinal image, and insofar as the retinal cells register the image, they also register the ellipse in some sense. So it appears as though this basic vision-scientific framework straightforwardly supports RET, and this might be why vision scientists often take RET for granted. Thus, the objector continues, RET might be a serious scientific claim after all, albeit a trivial one.

Response: the aforementioned retinal cells do not represent the ellipse *as* an ellipse, since this requires a mechanism for recognizing retinal shapes (the standard vision-scientific view is committed to the existence of mechanisms of this sort). The end product of such a mechanism, let us assume for simplicity, is a neuron that fires when and only when there is an ellipse of arbitrary size and location in the retinal image (this neuron in turn functions as input for processes that compute representations of objective properties). Now because whenever there is such an ellipse in the retinal image, in normal viewing conditions, there is also an object projecting it in the world, the aforementioned neuron fires, in normal viewing conditions, when and only when there is an external object with the relational property “having an elliptical projection” in view. Hence, the neuron tracks both the ellipse in the retinal image *and* the relational property of the external coin (its having an elliptical projection), which leaves the

question of what this neuron represents open. Thus, the fact that the visual system begins with the retinal image is not enough to support RET over RPV.^{4 5}

Here is the second consideration in favor of NSS. Hellie (2006) claims vision scientists are sometimes confused about the question of what perspectival properties are. As evidence, he provides examples of inconsistent claims about this matter made by Rock (1983). Additional evidence is that scientific literature reviews on the matter (his example is Todorovic 2002) do not detect a settled view on the question of what perspectival properties are. Hellie thinks that the scientists' confusion (or at least lack of consensus) arises because the introspective data bearing on this question is unclear. More specifically, he claims that introspection is silent about the nature of perspectival properties, specifically, it is silent on whether these properties are of external or of internal objects. This conclusion undermines the introspective (or phenomenological) case for RPV discussed earlier (recall the quote from Hill). In light of what I have said so far, it should be clear that there is an alternative, and more prosaic explanation for the scientists' inconsistent claims, namely that vision scientists do not seriously examine whether RET is true, making them adopt loose talk, leading sometimes to inconsistent talk, about perspectival properties. Thus, NSS, the claim that RET is not a serious scientific claim, neatly explains Hellie's data, which is a point in favor of NSS.

⁴ Some vision scientists (see. e.g., Todorovic 2002, p. 41) appear to be aware of this.

⁵ When confronted with this consideration (in personal communication), the vision scientist Shimon Ullman suggested that scientists' adoption of RET is merely a piece of *traditional scientific lingo*. He further predicted that, once confronted with the option of replacing (in their theory) representations of retinal shapes with representations of (suitable) relational properties of external object, most vision scientists might find it initially odd (because they are not used to talking this way), but they would not ultimately object to it.

This concludes my case for NSS. So, while RET causes trouble for RPV, RET is not a serious scientific claim, and so appealing to RET is not a good way to empirically test RPV. At the present point it might appear as though vision-scientific issues are orthogonal to the RPV/RET debate, making the debate a purely philosophical issue. But this is premature. In the next section I argue that a specific framework within vision science – that of object files – can be used to evaluate RPV.

3. Object files and RPV

The basic direction I want to develop here utilizes the influential framework of object files (Kahneman, Treisman, & Gibbs, 1992, henceforth KTG), according to which visual processing involve “object files”, which carry information about individual objects in the visual scene across some limited time interval. Object files involve two different kinds of representations and in two different ways: each file *stores* representations of properties, and each file as a whole *represents* (or refers to) an individual object. The file can be thought of as attributing properties to an object. KTG’s main evidence in support of the object files framework is a certain object benefit effect revealed in experiments. KTG’s experiments have roughly the following structure. A subject is required to name a certain letter on a screen, call it the target letter. It turns out that if the target letter matches a previously presented letter that is perceived as numerically the same as the target letter, the subject performs the task faster and more reliably, by comparison to a case where the target letter only matches a previously presented letter that is perceived as numerically different. KTG interpret the results as suggesting that information about shape (in this case, letter type) is stored in object files across time. When a subject is asked to recognize a feature of a target object, she automatically accesses, through the associated object file, information about the past shape of the target object, and this information helps her identify the current shape of the object.

Now if there are object files storing information about perspectival properties (henceforth "the perspectives in files" hypothesis, or PIF) *and* these object files represent external objects (henceforth EXT), then perspectival properties are attributed to external objects. This rules out RET, because according to RET, perspectival properties are properties of retinal regions, which are inner items. Remember that I am assuming the Error Theory is false. Given this, PIF plus EXT imply RPV. For, if the attribution of perspectival ellipticality to a tilted coin is veridical, perspectival ellipticality must be a relational property of the coin, a property such as "having an elliptical projection".

From the other direction, if RPV is true then perspectival shapes are attributed to external objects. Assuming with KTG that the visual system often organizes properties of objects into objects files of these objects, we expect perspectival properties to sometimes enter object files that represent external objects, in accordance with PIF plus EXT. Hence if PIF or EXT turn out to be false, this will count against RPV. At the very least, if PIF or EXT are false, proponents of RPV will need to explain why perspectival properties do not behave in the way expected in light of the object files framework. Thus, PIF plus EXT can be used to empirically evaluate RPV.

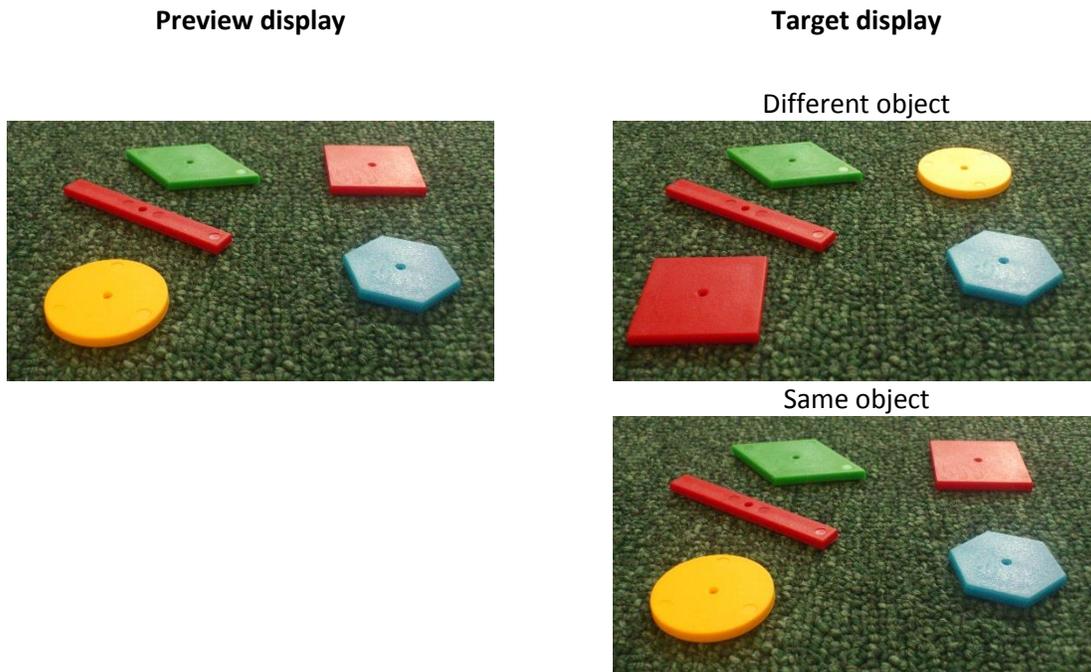


Figure 1. The two headings show two successive displays. Subjects are asked to identify the perspectival shape of the distant right object in the target display. In “different object” the target matches (with respect to perspectival shape) a preview object seen as a numerically different object. In “same object” the target object matches (with respect to perspectival shape) a preview object seen as numerically the same object.

What I would like to do now is describe in some detail vision-scientific considerations, theoretical as well as experimental, relevant to assessing PIF and EXT. Let us begin with PIF. One way to test PIF is by using KTG’s paradigm. That is, we need to look for the object benefit effect they have discovered, only now with respect to perspectival properties. The original KTG experiments, like many other classic experiments in vision science, do not involve a 3D scene; instead, they involve a flat, 2D scene played out on a computer screen. So it might seem that these experiments establish what we want, namely that information about perspectival properties is stored in object files. But this is premature. In the experiment setup, there are flat “objects” on a screen. Because in such a setup, the objects’ objective shape and perspectival shape are the same, say elliptical, one could interpret the results of the experiments in terms of objective shape alone, without mentioning perspectival shape at all. We thus

need to switch to an experimental setup where perspectival shapes and objective shapes of objects differ. We need a genuine 3D setup, in other words. Let me sketch an experimental setup of this sort. Keep in mind that this is only a sketch, which can be precisified in various ways. It is not meant to be an experiment that can actually be run as is. Consider figure 1. The experimenters present subjects with a target tilted object in 3D space, asking what its perspectival shape is, and then checking whether the fact that the object is seen as numerically the same as an object previously shown with the same perspectival shape facilitates recognition of the perspectival shape. If it does, we have the same object benefit effect described by KTG, meaning that perspectival properties enter object files (PIF is true). If it doesn't then this means that perspectival properties do not enter object files (PIF is false).⁶

⁶ Gordon and Irwin (1996, 2000) show that in KTG-style tasks of identifying letter type, differences in case or font between the target letter and the previously present letters do not reduce the object benefit effect KTG describe. On the basis of evidence of this sort Gordon and Irwin conclude that object files do not contain information about simple sensory properties, such as font or case (and hence shape); rather, they contain only more abstract information, such as information about letter type. Call this view ABS. If ABS is correct, then since perspectival properties are simple sensory properties, it follows that they do not enter object files, contrary to PIF. However, as Gordon and his colleagues admit (Gordon and Irwin 2000; Gordon, Vollmer and Frankl 2008), the evidence is compatible with hypotheses other than ABS. One is that the findings reflect task requirements, meaning that ABS is not true in general, but only with respect to tasks of identifying abstract features, such as letter type (Gordon and Irwin 2000, p. 149; Gordon, Vollmer and Frankl 2008, p. 677). This hypothesis is compatible with the possibility that in tasks of identifying perspectival properties, the latter enter object files. Hence the hypothesis is compatible with PIF. A second hypothesis is that in tasks of identifying abstract features, simple sensory properties enter object files, but they are not used to determine whether the target object is numerically the same as a previously seen one, which is why differences in simple sensory properties, between the target object and previously seen objects, do not reduce the object benefit effect, in the tasks in question (see Gordon, Vollmer and Frankl 2008, p.

Let me describe a different strategy for supporting PIF. It is sometimes thought that (a) object files mediate the conscious experience of persistence of objects (cf. Mitroff, Scholl, & Wynn 2005; Noles, Scholl, & Mitroff 2005), and that (b) when the visual system detects a new feature, it decides whether to put it in a given existing object file or to create a new object file for the feature, on the basis of the degree of dissimilarity between the detected feature and a feature *stored* in the existing object file (cf. Gordon and Irwin 1996, 2000). Taken together, these two claims imply that it is possible to study whether object files store features of a certain type by studying whether dissimilarities in features of this type influence the conscious perception of object persistence. If this is right, it is possible to support PIF by establishing empirically that dissimilarities between perspectival properties influence conscious perception of persisting objects. How might one go about establishing that? A natural starting point is studies examining the contribution of shape properties to conscious perception of persistence. I'll take a study by Feldman and Tremoulet (2006) as an example. In this study it is shown that, under some conditions, a certain degree of dissimilarity in shape properties between two objects, one approaching and disappearing behind an occluder, the other reappearing from behind it, makes observers consciously experience the two objects as numerically different. Like KTG's experiments, Feldman and Tremoulet's experimental setup consists of 2D displays, making the experimental results ambiguous between perspectival shape and objective shape. Hence their study, by itself, does not show that dissimilarities in perspectival shape influence conscious perception of object persistence. We've seen that it is relatively straightforward to make adjustments to KTG's experiments in order to avoid a similar ambiguity. Unfortunately, the same is not true with respect to Feldman and Tremoulet's experiments. Changing the displays to 3D displays in itself does not resolve the ambiguity, because changes in perspectival shape of an ordinary 3D object are always accompanied by changes in some other seen

677-8). This hypothesis is compatible with the possibility that perspectival properties enter object files even in tasks of identifying abstract features.

feature of the object, such as distance relative to the observer, degree of tilt, and (when non-rigid objects are involved) objective shape. It thus appears that the results of the modified study would be ambiguous between perspectival shapes and those other properties. Therefore, unlike the strategy of using KTG's paradigm in order to test PIF, the present strategy for testing PIF faces a seemingly difficult challenge, that of eliminating the said ambiguity.

Let us turn to EXT, the claim that object files storing information about perspectival properties represent external objects. There is an indirect way to argue for it (I discuss a direct way later on), via the hypothesis that object files with information about objective properties contain information about perspectival properties as well. Call this "the perspectival-objective object files" hypothesis. To deny this hypothesis is to hold that there are two sets of objects files, one for information about objective properties, the other for information about perspectival properties. If the perspectival-objective object files hypothesis is true, then since objects of files with information about objective properties are ordinary external objects, it follows that the objects of files containing information about perspectival properties are ordinary external objects as well, meaning that EXT is true.

Let me explain why I find the perspectival-objective object files hypothesis plausible. According to studies in vision science, there is an early visual process of grouping elements together, thereby segmenting the scene into objects, on the basis of representations of perspectival properties (see, e.g., Feldman 2007). For example (see figure 2), two lines that (perspectively) co-terminate tend to be grouped, and two lines that (perspectively) form a "T-junction" tend to be ungrouped. There is also an early visual process that solves, on the basis of perspectival properties, the "correspondence problem" for motion perception, namely the problem of determining whether two elements at different points in times are the same (Ullman 1979).



Figure 2. Example of a T-junction projected by two separate external objects, and of co-terminating lines projected by the same external object.

I will focus on segmentation, but similar considerations apply to the correspondence problem. In the object files framework, the said sort of segmentation is part of the process creating objects files (KTG, p. 210). Now, given that the scene is segmented early on into objects having perspectival properties, there is no need to segment it *from scratch* for the objective case. Arguably, then, objective segmentation (segmenting the scene into objects with objective properties) is (at least partially) *based* on perspectival segmentation (segmenting the scene into objects with perspectival properties). A plausible way to interpret this idea, given the object files framework and given the truth of PIF, is by suggesting that there are processes that create objects files containing – and on the basis of – representations of perspectival information, subsequently inserting into these files representations of objective properties, while updating the segmentations (i.e., splitting and merging files) if needed. Thus, each of the resulting object files contains information about both perspectival and objective properties, as the perspectival-objective object files hypothesis says.

The foregoing is not supposed to be an argument in favor of the perspectival-objective object files hypothesis. The point of the foregoing is merely to show that the hypothesis is a plausible interpretation

of the scientific data: the hypothesis can reasonably be used to flesh out the claim that objective segmentation is based on perspectival segmentation (henceforth “the basing claim”). The reason I say the foregoing does not by itself support the perspectival-objective object files hypothesis is that it is not mandatory to interpret the scientific data in the way I have suggested. It is in principle possible to flesh out the basing claim differently, in a way that does not imply the perspectival-objective object files hypothesis. Specifically, one might hold that although representations of perspectival properties help create and maintain object files containing information about objective properties, the former are not stored in the latter. Pylyshyn (2007, pp. 38-39) says something analogous to this with respect to representation of location. Specifically, he claims that although the mechanism of creating and maintaining object files makes use of representations of locations of external objects, these representations do not enter the object files. Perhaps the same is true with respect to perspectival properties.

One possible consideration in favor of the perspectival-objective object files hypothesis is parsimony. Why should the visual system create and maintain two sets of object files if it can do all the things it needs to do with one set? Creating and maintaining object files is costly. KTG (p. 208) show that the object benefit effect reduces if the number of objects in the scene increases, which they explain by suggesting it is costly to create and maintain these files.

A second strategy for supporting the perspectival-objective object files hypothesis is the following. There are direct ways of experimentally testing claims about whether two given representations are stored in the same file. Feldman (2007, p. 818), relying on results from Behrmann, Zemel and Mozer (1998), writes that “[c]omparisons of visual features are more rapid and accurate within a perceptual object than between distinct objects, a finding sometimes called the object benefit” (note that this kind of object benefit is different from the aforementioned one discussed by KTG). He uses this fact in order

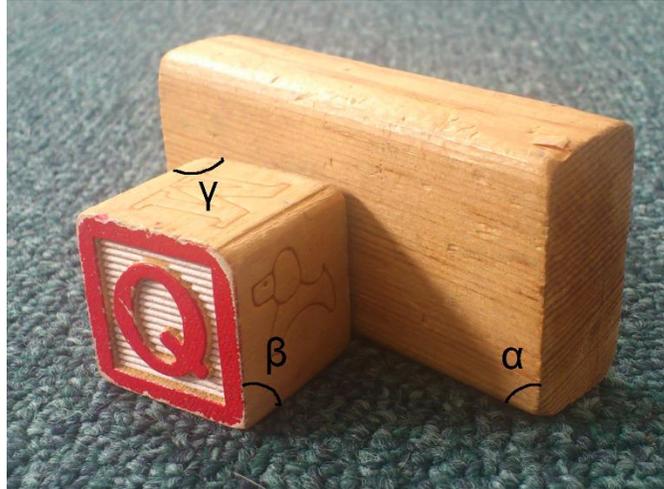


Figure 3. Subjects are required to assess, e.g., whether β understood objectively is bigger than α understood perspectively, or whether β understood objectively is bigger than γ understood perspectively.

to test whether (and how, and when) two given features belong in one perceptual object. Given the object files framework, the test determines whether two given representations of features are stored in the same object file. It looks promising to use this method to test whether (certain) perspectival and objective properties are stored in the same object file. Let me sketch an experiment of this sort. As the previous experiment I have suggested, this too is merely an outline of experiment, which can be precisified in various ways. Consider figure 3. Subjects look at boxes, and are asked to estimate the difference between the objective angle of a corner of a box and the perspectival angle of a different corner of *the same* box. Their second task is to estimate the difference between the objective angle of a corner of a certain box and the perspectival angle of a corner of a *different* box. If the perspectival-objective object files hypothesis is correct, subjects should be able to perform the first task quicker and more accurately than the second, because of the object benefit effect.

I now turn to explore a more direct way to argue in favor of EXT, a way that does not pass through the perspectival-objective object files hypothesis. The goal of the aforementioned early segmentation

process, which is based on perspectival properties, is to group pieces of information together in a way that corresponds to the way external objects are spread out in the perceived scene. To illustrate (see figure 2 once again), usually when the segmentation mechanism encounters a so called “T-junction”, each line is taken to belong to a different object (i.e., the lines are not grouped together), and this is the case because an objective situation that is likely to have a T-junction projection is one where one opaque external, ordinary object occludes another. Likewise, because co-terminating lines are typically projected by a single object, the visual system tends to group such lines together (for discussion of both examples, see Feldman 2007).

Here is what this means. Given there are object files storing information about perspectival properties (PIF), and given that the segmentation process groups together perspectival properties, it reasonable to hold that the segmentation process plays a significant role in determining, for any given file, representations of which perspectival properties it stores. We have seen that the segmentation corresponds to the way objects are spread out in the external scene. Thus, the resulting organization of information in object file also corresponds to the way objects are spread out in the external scene. Hence the objects that these files represent are ordinary external objects, not inner objects such as retinal regions. Thus EXT is true.

Conclusion

In this paper I have examined whether RPV can be tested empirically. In the first part of the paper I have argued against the claim that, because RPV conflicts with RET, and many vision scientists accepts RET, it follows that RPV straightforwardly conflicts with vision science. My argument was that while RET indeed conflicts with RPV, RET itself is not a serious scientific claim.

In the second part of the paper, I have suggested a novel, more complex way of empirically testing RPV. The way involves, first, checking empirically, on the basis of KTG’s paradigm or (more problematically)

on the basis of studies about conscious perception of persistence, whether information about perspectival properties is stored in object files (PIF). If it does, the second stage is to determine whether these object files represent external objects (EXT). They do so if and only if RPV is correct. I have explored a direct way and an indirect way to test EXT. The direct way checks whether the individuation conditions for files storing information about perspectival properties fit external objects. I have argued that there is existing data supporting an affirmative answer. The indirect way utilizes the hypothesis that files storing information about objective properties store (when appropriate) information about perspectival properties as well (the perspectival-objective object files hypothesis). I have explained that parsimony supports this hypothesis. Further, I have suggested that the object benefit effect utilized by Feldman (2007), which is different from the effect studied by KTG, could be used to empirically test the hypothesis.

Bibliography

- Behrmann, M., Zemel, R. S., & Mozer, M. C. (1998). Object-Based Attention and Occlusion: Evidence From Normal Participants and a Computational Model. *Journal of Experimental Psychology*, 24(4), 1011–1036.
- Biederman, I. (1987). Recognition-by-Components. *Psychological Review*, 94(2), 115–147.
- Brogaard, B. (2010). Strong Representationalism and Centered Contents. *Philosophical Studies*, 151, 373–392.
- Burge, T. (2010). *Origins of Objectivity*. Oxford: Oxford University Press.
- Burge, T. (2005). Disjunctivism and Perceptual Psychology. *Philosophical Topics*, 33(1), 1–78.
- Cohen, J. (2010). Perception and Computation. *Philosophical Issues*, 20, 96–124.
- Feldman, J. (2007). Formation of Visual “Objects” in the Early Computation of Spatial Relations. *Perception & Psychophysics*, 69(5), 816–827.
- Feldman, J., & Tremoulet, P. D. (2006). Individuation of Visual Objects over Time. *Cognition*, 99, 131–165.
- Fish, W. (2009). *Perception, Hallucination, and Illusion*. Oxford: Oxford University Press.

- Gordon, R. D., & Irwin, D. E. (1996). What's in an Object File? Evidence from Priming Studies. *Perception & Psychophysics*, *58*(8), 1260–1277.
- Gordon, R. D., & Irwin, D. E. (2000). The Role of Physical and Conceptual Properties in Preserving Object Continuity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*(1), 136–150.
- Gordon, R. D., Vollmer, S. D., & Frankl, M. L. (2008). Object Continuity and the Transsaccadic Representation of Form. *Perception & Psychophysics*, *70*(4), 667–679.
- Harman, G. (1990). The Intrinsic Quality of Experience. *Philosophical Perspectives*, *4*, 31–52.
- Hellie, B. (2006). Beyond Phenomenal Naiveté. *Philosophers' Imprint*, *6*(2). Retrieved from www.philosophersimprint.org/006002/
- Hill, C. (2009). *Consciousness*. Cambridge, England: Cambridge University Press.
- Kahneman, D., Treisman, A., & Gibbs, B. J. (1992). The Reviewing of Object Files: Object-Specific Integration of Information. *Cognitive Psychology*, *24*, 175–219.
- Lycan, W. G. (1996). *Consciousness and Experience*. Cambridge, MA: The MIT Press.
- Lycan, W. G. (forthcoming). Block and the Representation Theory of Sensory Qualities. In A. Pautz & D. Stoljar (Eds.), *Festschrift for Ned Block*. Cambridge, Mass.: MIT Press.
- Marr, D. (1982). *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*. San Francisco: Freeman Press.
- Mitroff, S. R., Scholl, B. J., & Wynn, K. (2005). The Relationship between Object Files and Conscious Perception. *Cognition*, *96*, 67–92.
- Noë, A. (2004). *Action in Perception*. Cambridge, Mass.: MIT Press.
- Noles, N. S., Scholl, B. J., & Mitroff, S. R. (2005). The Persistence of Object File Representations. *Perception & Psychophysics*, *67*(2), 324–334.
- Overgaard, S. (2010). On the Looks of Things. *Pacific Philosophical Quarterly*, *91*, 260–284.
- Palmer, S. E. (1999). *Vision Science: Photons to Phenomenology*. Cambridge, Mass.: MIT Press.
- Pylyshyn, Z. W. (2007). *Things and Places: How the Mind Connects with the World*. Cambridge, Mass.: The MIT Press.
- Rock, I. (1983). *The Logic of Perception*. Cambridge, MA: The MIT Press.

Schellenberg, S. (2008). The Situation-Dependency of Perception. *The Journal of Philosophy*, 105(2), 55–84.

Todorovic, D. (2002). Comparative Overview of Perception of Distal and Proximal Visual Attributes'. In D. Heyer & R. Mausfeld (Eds.), *Perception and the Physical World*. New York: Wiley.

Treisman, A. (1999). Solutions to the Binding Problem. *Neuron*, 24, 105–110.

Tye, M. (2002). *Consciousness, Color, and Content*. Cambridge, MA: The MIT Press.

Ullman, S. (1996). *High-Level Vision*. Cambridge, Mass.: The MIT Press.

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