# Objects, Seeing, and Object-Seeing

While driving, I see a car pull out into my lane: I hit the brakes. But if I had seen the shadow of a cloud cross my path, I would have just kept driving. I see a ball thrown in my direction. I raise my arm to catch it. It looks very different from a ball thrown toward the camera in a movie. The first member of each pair is an instance of what I call object-seeing: each is visual experience that elicits a reaction appropriate to a causally unified thing that I can physically interact with—a material object that can act on me as I can act on it.

### Three Questions

Here are three questions about seeing a material object.

1. *Seeing-as-such* What is it to see? Not merely to have an experience in which various properties are presented in a characteristically visual manner, but actually to *see*. What is the difference between *recalling* a car pull out into your lane—an episodic-memory experience in which the visual look of the event is repeated—and the experience of actually seeing it?
2. *Object-Seeing* What is it to see something *as an object*—not merely to be visually aware of a quality-instance or sense-datum that emanates from it? How, for instance, is seeing a car clearly enough to react appropriately different from peripherally seeing a brown blur that happens to be a car?
3. *Visual Particularity* What is it to see a *particular* object? What’s the difference between seeing one car and another visually identical one of the same make?

Most philosophical treatments of object-seeing focus on Questions 1 and 3.[[1]](#footnote-1) H. P. Grice’s (1961) Causal Theory of Perception (CTP), for example, successfully answers the question of *Visual Particularity*. But it is much less satisfactory when it comes to the other two questions. It gestures toward an analysis of *Seeing-as-such*, but wrongly (as I shall argue) shifts the responsibility for answering it to empirical science. And it is seriously misleading with regard to Question 2—it *evades* or *eliminates*, but does not explicate, object-seeing. My aim in this paper is to develop a better approach to the first two questions. (To the extent that he does not impinge on the first two questions, I will simply follow Grice with regard to the third.)

I propose a very simple approach to seeing-as-such. It is that veridical experience of things outside ourselves that is brought about by looking. *Looking* is something that we do, and of which we *can*, therefore, be self-aware. Thus, the experience of seeing is something that we are able, in the right circumstances, to recognize by its phenomenology. (Grice famously demurred, arguing that an intervention from empirical science was needed to define seeing—as I said earlier, he shifted responsibility.)

As to the question of object-seeing, material objects appear as the bearers of visual features. A material object does not take its place in the visual field alongside features such as colours, shapes, and movements. Rather, it appears as the bearer of such features—a concrete body that is colour*ed*, shap*ed*, mov*ing*. It appears, moreover, as a massive object with which one can interact. Mere located features do not merit physical preparations; one doesn’t brake for or attempt to catch what appears to be a disembodied colour, only for a material body.

As we’ll see, Grice’s CTP attempts to account for object-seeing solely in terms of a mind-independent causal relation that does not impact on phenomenology. Thus, it fails to explain the special look of material objects. So, it leaves us in the dark about the origins of the reactions mentioned just now. To account for these, I appeal to a distinctive phenomenology of object-seeing. I will show how certain other visual experiences, like seeing a face in a mirror or a person in a picture, fall short.

Phenomenology-lite versions of CTP have led a number of philosophers to counter-intuitive claims about visual objects. Helen Yetter-Chappell (2018) claims that you see your own face in a mirror in exactly the same way as you see the face of your companion across a table. Roy Sorensen (2008) says that we see the rearward facing sides of back-lit objects. Kendall Walton (1984) contends that we see absent things in photographs (but not in paintings). Taken as claims about the external causal origins of visual impressions, these claims are provocative and surprising, but ultimately innocuous. Taken as claims about object-seeing, they miss the mark. I will show how..

### What is a Visual Object-Impression?

Grice writes:

for an object to be perceived by *X*, it is sufficient that it should be causally involved in the generation of some sense-impression by *X* in the kind of way in which, for example, when I look at my hand in a good light, it is causally responsible for its looking to me as if there is a hand before me, or in which . . . (and so on), *whatever that kind of way may be*; and to be enlightened on that question one must have recourse to the specialist” (*ibid*,143-44; the ellipsis is in Grice’s text).

The claim is that I see an object only if it is a cause, in the right way, of some *visual impression* in me. Central examples of *hearing* and *touching* would, no doubt, take their place among the ones elided by the “or in which . . .”, but for present purposes, I’ll mainly consider *seeing* external objects.

Following H. H. Price’s (1932) lead, Grice’s CTP is designed to elucidate object-seeing in the context of sense-datum theory. Sense-datum theory is highly reductive. In it, we only really see quality-instances in visual field-places (“Blue there”). So, it leaves a puzzle. We commonly say that we see material objects: people, faces, cars. What do we mean?

There are two approaches to this question consistent with retaining sense-datum theory. The first is a *deflationary* answer about object-seeing—plausible truth-conditions for statements about material objects that do not require any modification of the theory. The second approach is *constructionist*: itsays that we do indeed have visual experience of material objects *over and above* (mere) sense-data, but reduces these objects to complex constructions out of sense-data.

Price, and following him Grice, take the deflationary approach. For them, ‘*X* (an object) looks blue to me’ means simply that (a) I experience a blue sense-datum, and (b) my experience is caused by *X*. This approach requires nothing over and above awareness of atomic sense-data. The blue-impression does not have to be a constituent of a bigger construction; it doesn’t even have to be visually separate. For on this account, you could say that you see a successfully camouflaged moth on the bark of a tree—it is, after all, responsible for a bark-like sense-datum. But there is a clear sense in which you *don’t* see the moth in these circumstances—a child playing hide-and-seek couldn’t say “I see you” in this sense. I’ll return to this point in section IX below.

More centrally to my purposes, you could, as far as Price and Grice are concerned, have a non-object-like impression of *X*, a mere flash or blur that doesn’t distinguish it as an object—it does not appear as a persisting subject for other features and does not elicit any object-appropriate reaction. This would count for them as seeing the object. This is a form of deflationary reduction. Their treatment preserves sense-datum theory—but at the cost of failing to acknowledge that one sees objects *sensu stricto*. It simply does not discern any role for objects in visual awareness other than as external causes.

This is a serious shortcoming, for not only does CTP fail to do justice to an important distinction within visual phenomenology, but it is also unable to account for an important difference between vision and certain non-visual modalities. Visual experience of objects in direct and unmediated, whereas in audition, for instance, experience of material objects is mediated by experience of sounds. CTP has no resources to account for this difference; as far as it goes, we hear material objects in precisely the same way as we hear them.

### Visual Particularity

As I said before, Grice’s main goal was to specify *which* object someone sees—the question of particularity. To take his now familiar example, you are standing in front of a mirror in a hall of pillars. You see a pillar reflected in the mirror, and there happens to be an exactly similar pillar that occupies the virtual location of the mirror-image. Which of the two pillars do you see—the one behind the mirror or the one behind you? Your visual impression is exactly as it would have been if the mirror hadn’t been there. But the pillar behind the mirror is “causally irrelevant” to your visual impression (142)—it is the reflected pillar that is causally responsible. Thus, according to CTP, you see the reflected pillar that happens to be behind you, not the pillar behind the mirror.

Despite his failure to address object-seeing proper, Grice was at least partially correct in his answer to the question of particularity. Whether you know it or not, your visual experience reveals something behind you, not something behind the mirror. But we have seen that sense-datum theory cannot account for what it is to see a *pillar*, which is a kind of material object, as distinct from having a vague visual impression of a blurred smudge of colour. By the same token, it lacks the resources to address the question whether you *object*-see a reflection of yourself *in* a mirror. I’ll argue (section XII) that there are grounds for doubting this.

### Seeing-as-Such

Even more basic: what is a *visual* impression? Grice says that it is a state of a perceiver reported by:

. . . such supposedly standard locutions as “So-and-so looks *Φ* (*e.g.,* blue) to me”, “It looks (feels) to me as if there were a *Φ* so-and-so”, “I seem to see something *Φ*” and so on. (1961, 123)

But this is plainly insufficient. For as Grice acknowledges almost immediately after the above quotation (only to push the problem offstage), words like ‘looks’ have non-visual applications. Consider this example:

(1) To me, his finances look dangerously compromised.

Here, the bare causal condition may well be satisfied: the man’s finances may be causally responsible for the reported impression. However, it’s obvious this is not a *visual* impression.[[2]](#footnote-2)

This vitiates CTP. Grice had an avenue of inquiry open to him—clearly, the man’s finances are not causally responsible for my impression in the way that my hand is responsible for my visual impression of it. That is, the finances do not create an impression on me by sending an optical signal to my retinas. But he does not pursue this avenue of inquiry. Doubtless, this was because he thought that when we discuss “optical signals,” “retinas,” and the like, we stand in need of “recourse to the specialist” because such talk is outside the philosopher’s domain. However this might be, he rests with his vague allusion to standard uses of the “looks to me” locution. CTP does not even get going until the notion of a visual impression has been clarified, but Grice does not undertake the task.[[3]](#footnote-3) I’ll turn to it now.

### Setting Semantics Aside

Some attempt to use semantic analysis to address the over-broadness of the ‘looks’ locution. They say that the sense of ‘looks’ in example (1) above indicates a cautiously held belief, not a visual impression. Roderick Chisholm (1957, Chapter 4) and Fred Dretske (1969) take this view; they say that there is a particular sense of ‘looks’ that reports visual experience. Chisholm calls this the ‘non-comparative’ sense, Dretske called it “non-epistemic.”[[4]](#footnote-4) Since I am going to come back to Dretske in my consideration of *object*-impressions, I’ll adapt his version of the claim:

‘*M* looks a certain way to *S*’is true in the *non-epistemic* sense of ‘looks’ if this does not analytically imply, for any proposition *p*, that *S* believes *p.*[[5]](#footnote-5)

Supplementing Grice accordingly, we have:

*Grice-Dretske Semantic Definition* A visual impression is (by definition) a state of a perceiver reported by locutions such as ‘So-and-so looks *Φ* to *S*’ in which ‘looks’ is used non-epistemically.[[6]](#footnote-6)

The Grice-Dretske definition has some initial plausibility. Contrast (1) with

(2) Those shoes look black to me.

The claim is that the only way to understand (1) is as attributing a belief to me. If that belief changes, (1) becomes false. (2), on the other hand, reports a state that need not be connected with belief—I can change all of my beliefs and it could still remain true. Visual states, then, are those that can be reported using the word ‘looks’ without implying anything about the subject’s beliefs.

But now consider:

(3) Those shoes look expensive to me.

On the face of it, (3) reports a state of seeing-as-such—a more complex state than (2), to be sure, for it involves more than colour—but a visual state that is, for this very reason, reportable only in this indirect manner. But (3) implies a belief about how expensive shoes look.

A defender of Dretske’s approach could respond to this challenge by saying that the visual state in question is more properly fixed in some other way. Frank Jackson (1977) suggests this route—he says that there is a special sense of ‘looks *F*’ that takes colour, shape, or distance for *F*. But this raises the question whether every visual state necessarily reduces to these low-level visual properties. After all, there are many holistic visual looks—the look of an early spring day, the look of a Dutch master painting, the look of man who is drunk, and so on.[[7]](#footnote-7) One can’t reduce these recognitional looks to low-level visual configurations. (3) reports just such a holistic visual impression; many of these are belief-dependent.

A related point: The look of expensive shoes might be *affectively* influenced by belief: that is, it could be influenced by beliefs about the virtues and vices of spending a lot of money on a pair of shoes. The shoes might, for example, *look* more desirable, or possibly more repulsive. Various views of “top-down” influences on perception (Hansen et al 2006, Balcetis and Dunning 2010, Macpherson 2012, Stokes 2012) imply that this is a real possibility. All of this opens up the possibility that there could be an affective component of visual perception that is connected to belief. Thus, there might be visual perceptions that are best reported epistemically (in Dretske’s sense). Without a substantive (and to my mind, implausible) view about visual impressions, one cannot rule out this out. It’s wrong to *define* visual perception in a way that depends on a contingent and contested assumption about it.

### The Phenomenology of Seeing

There’s a better approach.

In a different context, Grice (1962) asserts that seeing has a “special introspectable character”—a characteristic phenomenology—that distinguishes it from perceiving with the other senses. One could therefore side-line the idea that visual states are identified by a characteristic use of the verbs, ‘look’ or ‘see.’ Instead, one could simply say that one sees *X* if *X* is causally responsible for an experience of things outside oneself that is phenomenologically characteristic of seeing.[[8]](#footnote-8) With this in mind, let’s move to the phenomenology of visual impressions.

Now, it’s a common assumption among philosophers of perception that the “special introspectable character” of vision (SICV) is to be found in the *content* of visual experience, specifically in the properties to which vision gives us access. Grice held this view: he thought that SICV could be defined by generalizing from particular experiences of visual properties. Consider what it means to say that we see colour. The specific colours are *red*, *blue*, and so on; each of these is associated with a specific kind of experience, *looks red*, *looks blue*,and so on. *Colour* is the genus of the specific colour properties and *looks coloured* is the genus of the experiences common to all of the colour looks. In the same way, he says, SICVis what is shared by all of the experiences marked by *looks F* for values of *F* that span the characteristically visual properties.

in addition to the specific differences between visual experiences, signalized by the various property-words employed, there is a generic resemblance signalized by the use of the word ‘look,’ which differentiates visual from nonvisual sense-experience.

Grice once again introduces the word, ‘look,’ but as noted earlier, he is bracketing the difficulties we have been trying to deal with. His proposal is that there is something common to all experiences of looking as if something has a visual property. This commonality defines what it is to be a visual experience.

Now, Grice’s proposal fails as a way of defining the phenomenology of *seeing*. Visually imagine an expansive field of green grass. This imaginative experience is phenomenologically very different from actually *seeing* green grass. But this difference cannot be captured by the look of visual properties; you visually imagine the grass by mentally recreating the same visual look as the grass that you see. (This is why Hume held that the green you imagine was qualitatively the same as the one you see, though less “vivid.”)

Defenders of Grice might say that this is precisely where CTP comes in handy: *imagining* a field of green grass does not count as *seeing* it because the imagining is not caused by a grassy field in the right kind of way. True: but my point is that the phenomenology of *seeing* green is *intrinsically* different from that of *imagining* green—it is phenomenologically different—and not just because of its causal antecedents. I believe that I *see*, rather than *imagine*, a field of grass because my experience is characteristic of seeing, as opposed to imagining. I don’t check on the causal antecedents of the experience to make this judgement.

Now, at first glance, John Searle’s (1983) account of seeing might seem like a way out of this difficulty. Importing Grice’s CTP into visual content, Searle proposes that to seem to see a material object *M* is not just for *M* to look as if it has certain visual properties, but additionally for it to look as if *M* were causally responsible for that very experience of seeing. But this doesn’t help. As has often been remarked, visually to imagine an object is to imagine how it would be to see it. So, visually to imagine a material object *M* is (if Searle is right) to imagine its looking as if it is causally responsible for the visual impression it produces. You can’t distinguish *seeing* *M* from *imagining* *that you see M* by dwelling on how *M* looks when you see it—that same look is repeated in the imagination. (*Mutatis mutandis* the same goes for recollection.) To put it briefly: seeing is not distinguished from recalling or imagining by the *content* of these experiences.[[9]](#footnote-9)

Despite the above, it is clear that seeing *does* have a characteristic introspectable character; you seem to be *seeing*, as opposed to imagining or recollecting. So even if SICV is partially characterized by visual content (however that is understood) there must be some residue that distinguishes *seeing* from other visual experiences. My suggestion is that this residue lies in the activity of *looking* at a thing. *Looking* at something is *phenomenologically* different from imagining or recollecting it. *Seeing* an object is a state produced by the activity of looking, while imagining it or recalling it produces a similar image by means of a different activity. When you look at anything, you have an experience that appears to be produced by this activity. Similarly, when you imagine or recollect anything, you have an experience that appears to be produced by the activity of imagining or recollecting. These activities are phenomenologically different, and you are normally self-aware of which one you are engaged in.[[10]](#footnote-10)

To aid in elucidating this idea, here are two perceptual activities that phenomenologically distinguish looking both from visual imaging and recollection, on the one hand, and from listening, smelling, touching, etc. on the other.

1. *Using the eyes* The impressions produced by looking seem to result from using my eyes. They go away when I close my eyes; they change when I squint or squeeze my eyes, or place a coloured filter in front of them. By contrast, my eyes are self-evidently not at work in my visual imaginings. Seeing, in short, is a state that is produced by looking; seeming to see is a state that is produced by seeming to look.
2. *Using the body* When the observer moves, there are characteristic systematic changes in experience that correlate with changing distance, direction of illumination, and within-object coordination. These correlations can be used to gain information that is absent in a static image. The way an observer moves to gain this kind of information is part of the perceptual strategy that she employs.[[11]](#footnote-11) She turns her head and moves her eyes; she manipulates the object and changes the lighting, e.g., by drawing the curtains and letting the sunlight in. Such changes of viewing circumstance are voluntary and free; we gain information from how the visual experience changes in response to these changes of perspective we cannot gain this information from a static, monocular point of view.

These interactions between perceiving subject and perceived object are constitutive of the appearance that the perceived object is independent of the perceiver (Siegel 2006) and causally responsible for the perceiver’s visual impression. For example, when I move, the object presents a different side to me—if it doesn’t, and if it is close by, then it must be a figment. Thus, these explorations—active looking—constitute the phenomenology of seeing, as opposed to imagining or recalling.

Now, it is, of course, possible that I might experience the phenomenology of looking, when in fact I am not looking—perhaps this could happen in an especially vivid dream, for example. Conversely, it might appear to me that I am merely imagining a faint light in my visual periphery where there is in fact a light that I really see. And Grice was particularly concerned about what David Lewis (1980) called “prosthetic vision”: veridical visual-like experiences that are produced extrinsically—for instance, in a sighted person by direct stimulation of the visual cortex, or in a blind person by sensory substitution.[[12]](#footnote-12)

These are legitimate concerns, but I think that Grice took a wrong turn in addressing them. He wanted to say that the difference between such cases and actual vision lay in the nature of the process—for example, in the fact that actually seeing an object occurs by light from that object stimulating the retina. And he thought that defining this process was a matter that fell outside the purview of philosophy: to determine how actual vision is special requires “recourse to the specialist.” I think this is wrong. We don’t need to know anything about the retina to know whether these cases count as seeing; what we need to know is whether the visual states are produced by looking. True, we can’t know by mere introspection whether a visual state is produced by looking, but specialists are not needed to specify what looking is.

Summing this up, we have:

*Seeing* is veridical experience of things outside the subject that is produced by *looking*.

Following Hume, I will say that you experience *visual impressions* when you see. Similar definitions can be given of auditory, tactile, etc. impressions—in each modality, perception is the product of a characteristic activity (Matthen 2015b).

This solves the problem of shared content. Seeing is phenomenologically different from visual imagining and visual recollecting, and also from hearing, tasting, feeling, and smelling, not because its *content* is different, but because it appears to be the result of looking.

### Point-Content vs Object-Content

I’ll turn now to my second question—What is object-seeing?

Some sensory states have mere *point-content*. That is, their content is about a single undifferentiated location in space and time. Think of a sniffer dog following a trail. As long as the odour gets stronger or stays the same, she continues along her path. But if, as she sniffs, the odour weakens, she backs up and tries another direction. The dog is responsive to a property of an odour *field*, or distribution of odour strength in space; she is following a gradient in this field. Each of her olfactory states is of the form: <*Q*, here, now>, where *Q* is an olfactory quality (which implicitly includes the strength of the smell). Her behaviour is guided by the comparison of such states over time. In the terminology I just introduced, the olfactory information she requires is merely point-content.

Now, think of logical atomist accounts of the visual perception of shape. The claim is that in an analogous way, this is just a spatial aggregation of point-content. What we see is a ‘colour mosaic’—the term is due to Lewis (1966)—an array of coloured points across the visual field. To see a circle or a square, it is argued, is merely to see a contiguous collection of points that collectively happen to stand in that particular spatial pattern. In the terminology of section II, above, this is a constructionist approach to shape (and object) seeing.

Here is the crucial point. As *active agents*, animals are interested in more than point-content and its spatial distribution. They are interested in potential mates and predators, sources of food, moving objects that they must avoid or track according to interest, and landmarks that help them navigate their surroundings. These things are *material objects*—connected, causally coherent, spatiotemporally extended, continuous collections of matter. When you perceive a material object, you don’t merely seem to see a collection of visual-field points in which there are sensory features. Rather, your visual system delivers to you the phenomenological appearance of an entity that, by virtue of internal causal inter-relations, coheres as a single unit.

Here are some of the ways that material objects are visually different from collections of visual field points:

*Logical subject* Material objects look as if they are the bearers of visual features. When they move, they usually take their features with them, and when they do not, they look as if they are qualitatively changing. (Pylyshyn 1989 is a canonical study of visual objects that persist through change and movement.) We may take from this that in object-seeing, we do not merely see features in a spatial location, *L*, but a feature that belongs to an object, *O*, which is (for now) in *L*.

*Spatial location* Material objects look as if they occupy a definite location in *space*; they don’t just seem to occupy a place in the visual field. Thus, there is a phenomenal difference between how things look when an object moves and the subjects stays still and when the subject moves and the object stays still. Moreover, material objects visually appear as if their visual field position would change in more or less predictable ways if the subject were to move (Siegel 2006). Reflections and depictions lack this characteristic—they do not appear to occupy a real place. Mirror images behave anomalously as you approach them; their distance from you decreases at twice the pace of seen material objects. Depicted objects do not appear to be located in the space that you occupy; they are not visually represented as being a definite distance away from you (Matthen 2005, chapter 13). This marks reflections and depictions as not being denizens of external space. They do not look as if they would resist the intrusion of other objects.

*Availability for interaction* Material objects look as if they can be grasped or touched. This is not true of depicted objects, two- and three-dimensional cast shadows, beams and flashes of light, fogs and mists, and holes.

*Amodal Completion* Material objects look as if they possess unseen rearward facing surfaces. This is not true of two-dimensional shadows or depicted objects or mists. It’s not obviously true of beams of light (though it is true of transparent objects).

*Interpenetration* It is possible for one material object to be seen as penetrating another, as when a knife cuts butter. Contrast a beam of sunlight coming through a window and falling on a table. There is no visual impression of interaction between the two.

*Unitary action* Material objects cohere in complicated ways. When a dog eats from a bowl, the bobbing movement of its head communicates itself to its haunches. It matters to us when we encounter such causally integrated wholes, and our perceptual apparatus is organized around detecting them.

The above are some of the phenomenological difference between the impression of aggregated point-content and that of a material object.

### Non-Visual “Objects”

The best evidence that animal perception treats material objects as a special focus is this: There is no precise analogue of visual point-content in the non-visual modalities; yet there are object-analogues: sounds, smells, flavours (O’Callaghan 2007, 2016, Nudds 2010, Matthen 2010, Fulkerson 2014, Batty 2015, Smith 2015). And these non-visual objects are normally organized around material objects. This shows that material objects have ecological significance that transcends the mere aggregation of visual point-content.

Let’s look at how object-analogues are extracted from raw input in non-visual modalities.

* The input to hearing is the vibration of the air at each ear. From this, the auditory system extracts individual voices, phonemes, melodies, and other such coherently structured acoustic entities. These are *sounds*.
* The input to smelling is a mass of molecules; from this, the olfactory system extracts a number of *scents* or *smells*, each of which consists of more than one kind of molecule. These smells pertain to individual material objects from which they emanate.
* The input to touch is somewhat like visual content in that it is place-wise distributed. However, this input is indexed not by external spatial location, as vision is, but by place on the body. Importantly, a stimulus on, say, the finger seems to come from the same place even when the finger is felt to move. From these stimuli, touch extracts the shapes and feels of material objects.

As an index of the primacy of material objects, consider Roberto Casati’s (2015) observation:

Preference for whole objects is evident in perception and action from early infancy . . . In the absence of clear contextual indications to the contrary, new names are by default assigned to an object semantics rather than being taken to refer to any other type of entity (*ibid* 402)

Think, then, of two contrasting pairs of scenarios:

1. (a) You and I walk into a garbage dump and are assailed by an overwhelming stink. I say, “That’s disgusting.” Clearly, I am not referring to the smell of any particular object.

(b) You are sniffing a glass of wine. I say “That’s sauvignon blanc.” Here, it’s perceptually obvious I am referring to a single complex odour that comes from the liquid in the glass.

Casati’s point is that the “default” is to assign the name to an entity organized around a material object, where such an object is evident. In both cases above, your nose picks up a large collection of different odour-molecules. However, in (b), but not in (a), the olfactory sense combines them into a unitary scent that it attaches to an object that you can identify by vision and touch.

2. (a) You’re walking in the woods and a wind springs up creating a rustling noise all around you. I say, ‘That’s a Chinook.’ Even though you are not familiar with this name for the wind, you may assume that I am talking about the totality of sound, not the particular creaks and groans coming from individual trees.

(b) You are listening to a symphony. Suddenly, on the right a reedy new sound enters. I say, “That’s an English horn.” It’s evident I mean a sound coming from a particular instrument, not the sound of the entire orchestra.

In both cases, your ear picks up a composite of frequencies, but in the second it is able to identify a coherent (“reedy”) sound that emanates from an individual object that you can visually locate. The difference, as before, is that a non-visual system is able to unitize a particular collection of frequencies and recognize a coherent sound that it can attach to a visually identified object.

The difference between the (a) and (b) members of these pairs illustrates how the unitizing phenomenology of the non-visual senses is, wherever possible, organized around material objects, with regard to which, in turn, visual object-perception takes the lead.

The scenarios above illustrate that there are separate levels of organization at work in each modality: olfaction unitizes odours; audition unitizes voices and speech streams, touch unitizes temporal sequences of complex tactile sensations, and so on. Perceptual systems construct these non-visual units because of correlated patterns that they find in the sensory stream—patterns that indicate a single source that is identified visually.

As Casati suggests, this is the default case. However, there are exceptions. Sometimes, there is no visual object that the non-visual unit can be attached to. This is what happens in the (a) cases above. And in these cases, sounds and smells tend to lose their unitary character, being experience, instead, as a melange. On other occasions, the non-visual unit is produced by the coordinated activity of several visually distinct material objects, which then sound as if they were one—the sound of an orchestra is an example of this.

### Visual Segregation

Grice’s Causal Theory of Perception does not demand that when we see an object, we should visually be able to discern it as an object. It focuses on point-data and fails to acknowledge object-content. Thus, as noted earlier, it fails to exclude a familiar sort of case:

*Camouflage* There is a perfectly camouflaged animal in front of you—say a moth that perfectly blends in with the tree trunk on which it lies. No matter how closely you look, and no matter how you adjust your perspective, you simply cannot visually discern the presence of the moth.

Do you see the moth? On one perfectly normal understanding of the question, the answer is ‘no.’ In good conditions, vision demarcates an object, but here it doesn’t reveal the moth’s presence. This case is somewhat similar to failing a test of colour vision: a red-green colour-blind person shown an Ishihara plate with the numeral ‘5’ written in orange dots against a uniform field of green dots. She simply sees a uniform field of dots.[[13]](#footnote-13) She doesn’t see any number written there. Similarly, in the *Camouflage* case, the visual process of object-contour recognition fails; you just don’t see the moth.

I said earlier that Grice is *not* concerned with the impression as of a distinct material object as from the impression merely of a quality-instance that can be attributed to an object. Because of this, he was not concerned that on his definition, you see the camouflaged moth.[[14]](#footnote-14) The moth is a patterned brown. Light reflected off the creature reaches the eye and successfully engages colour vision; the patterns on the moth’s wings can be discerned (though not differentiated from those of the surrounding bark). Presumably, the moth is causally responsible for this impression in just the way that a hand in good light creates a visual impression in you. This is proved by the fact that you *can* see the various markings on the moth’s wings. It’s just that since the moth is camouflaged, these look as if they belong to the tree on which the moth is sitting.

Dretske’s positive account of seeing is designed to accommodate the intuition that you *don’t* see the moth. According to him, you see an object if and only if you visually differentiate it from its immediate environment. By this, he means that there is some visual characteristic or mark that enables you to distinguish the object from its surroundings. The camouflaged moth is not distinguished from its surroundings by any visual mark. So, you don’t see it.

Dretske’s account of object-seeing is an advance over Grice’s because it acknowledges visual segregation. But it is insufficient to capture object-seeing proper. Visual segregation is consists merely of cordoning off a part of the visual field that is qualitatively different from its surroundings. However, mere field-regions do not as such exemplify the characteristics of material objects listed in section IV above—logical subjecthood, spatial location, availability for interaction, amodal completion, etc. For instance, visual objects such as shadows, reflections, and even migraine auras,[[15]](#footnote-15) appear to be visually segregated—they are associated with regions of the visual field—but they do not, or do not fully, have the appearance of material objects. Migraine sufferers see auras, and Dretske’s criterion accommodates this, but they do not see object-see them. Thus, as I will elaborate in what follows, visually segregating a region of space is neither sufficient nor necessary for seeing material objects as material objects.

A brief remark that I do not have the space to elaborate fully here. Material objects are more fundamental in object-vision than other substance categories because they are logical subjects. As well, as I argued in Section V, visually identified material objects also serve as perceptual anchors for non-visual units such as sounds. Material objects are privileged in this kind of perceptual organization: there is no more specific configural organizational kind. Think of the more specific sortals—*animal*, *tree*, *rock*, etc. Vision does not have special principles for these. To put this in another way, tree-vision, animal-vision, etc. do not mark off distinct visual phenomenologies in the way that object-vision does. *Material object* is the only visual sortal. (For further discussion, see Xu 1997 and Matthen 2005, chapter 12.)

### Reductive Approaches to Object-Perception

A great deal of the scientific study of object perception focuses on how object-content is constructed from point-content. And the detection of object boundaries, or visual segregation, is often an important part of this construction. Two birds flying together display many correlations and appear, in many ways, to be a single perceptual unit. Yet they are seen as *two* birds, while each of them is seen as *one*.Why? Perhaps, because each is visually segregated from the other. But this kind of example should not lead us to think that object-seeing *reduces to* the detection of object boundaries. Visually recognizing something as possessing spatial boundaries need not involve seeing the boundaries themselves. In the case of the proverbial dog-behind-the-fence, you see the dog without seeing its outline. Conversely, seeing boundaries may not entail seeing an object contained within the boundaries. Seeing a dog chasing a ball is very different from that of seeing a *movie* of a dog chasing a ball. When you watch the movie, you see patterns of light and shadow that mimic the visual segregation of a dog—but the moving image of the dog does not look as if it is a material object that occupies a definite three-dimensional region of space. The problem of how certain regions of the visual field get segregated as single units—the problem of spatial unitization, let us call it—is different from that of how we detect material objects.

To make this point more concretely, let me take up a recent putative account of “perceptual objects.” E.J. Green (forthcoming) proposes a highly sophisticated reductive approach to what he takes to be the phenomenon of object perception—though in reality it is an approach to spatial unitization. Very schematically, his proposal is that a perceptual “object”—I’ll call *his* objects “units” to avoid confusion—is a collection of “parts” that display causally sustained perceptual regularities among themselves, such that the addition or subtraction of candidate parts reduces the degree of regularity manifested.

Green’s approach affords us a revealing and innovative account of how spatial units are constructed in vision, and also in audition.[[16]](#footnote-16) However, Green believes that the function of spatial unitization is coding efficiency. Perceptual systems gather point-data about surrounding spatial regions, with the goal of providing relevant information to psychological systems involved in action and belief-formation. The problem, according to him, is that the perceptual data-set is dauntingly large and needs to be compressed before it can be transmitted to other systems. Perceptual systems employ regularities within the data-set to compress it. Objects, as he sees them, *emerge* from these coding efficiencies; they are not the primary focus of functional interest.

Green maintains, then, that material objects are *not* (except incidentally) the targets of spatial unitization. But in arguing for this negative conclusion, he commits what I take to be a fundamental but instructive error: a conflation of volumes with the things that occupy them. Green takes on the problem of explaining how we see things like apples and geese, which we differentiate from their surroundings. But he takes as his target an antagonist “body theory” of how we perceive these things. In this theory, “bodies” are defined topologically: they are coherent, bounded, continuous, maximal, three-dimensional volumes. But, he says, spatial unitization does not, in fact, single out only bodies understood in this way.

The body principles are obviously more restrictive than the perceptual organization principles. Certain collections, like flocks of geese, are noncohesive but perceptually groupable. Certain undetached parts, like arms, are nonbounded but perceptually parsable. These entities certainly seem to constitute “units” in visual experience. But the body theory says they aren’t visual objects.

So, he concludes, the body theory fails.

I agree with Green’s rejection of the body theory. But I don’t agree with his reasons. Contrary to his view, we *don’t* see flocks of geese as material objects: we would not be surprised if a crow were to fly right through them; our visual expectations are not violated when they split apart. Similarly, we don’t treat undetached parts as if they were material objects: we are not afraid that shaking a man’s hand will result in it becoming detached. Thus, I want to say, the body theory shares with Green’s own a much more fundamental misconception: both suggest that object-seeing is nothing more than the visual segregation of spatial volumes, and thus it fails to distinguish these volumes from the visibly distinct objects that occupy *some* (or rather, many) of them.

Suppose, then, that you see a green circle that appears to move across the visual field without changing shape or colour. This is a visual unit on any account. The point of disagreement is how the unity should be defined. According to body theory, the moving green circle is merely a topological continuity in temporally successive visual presentations. According to Green, it is a maximal configural regularity in point-data—the visual system uses this regularity to compress the data-set. According to me, it is a material object; it looks like the logical subject of the visual features, *green*, *circular*, *moving*, etc. It is, moreover, the locus of possible physical interaction, and it looks as if it has a rearward-facing side. This is a disagreement about both the phenomenology and the ecological significance of perceptual unitization. Green suggests that one comes to see material objects accidentally, as a by-product of a data-compression routine. And he does not acknowledge the additional phenomenological accoutrements of object-seeing. On the basis of the principles outlined in section IV above—logical subjecthood, spatial location, availability for interaction, amodal completion, etc.—I would argue that material objects are the target of object-vision.

The phenomenon of visual objects is not the same as that of spatial unitization. You see lots of spatially bounded things; some look like material objects, some don’t. A field of tall grass undulating in the wind is perceptually segregated; it does not look like a material object. A dog normally looks like a material object. It is a three-dimensional body that does things and has things happen to it as a single unit—and that’s the way it looks. A *movie* of a dog is a visibly segregated two-dimensional patch of colour that incorporates many of visual coherencies that the dog presents. (That’s why it’s a picture of a dog.) But it does not look like a material object; it does not look as if it is doing anything. (If you hold your finger in front of the projector, a shadow falls on the screen; it looks completely different from a shadow falling on a dog.)

The Gestalt psychologists looked to two-dimensional displays to tease out some of the principles of visual material-object identification—principles of continuation, occlusion, and so on. They sometimes suggest that object-perception is simply a matter of boundary construction or visual segregation. This is what Dretske also suggests. This is the origin of body theory. But it is possible to allow that visual segregation is one way, perhaps the most important way, we identify bodies without saying that object-perception is just a side-effect of visual segregation. The discussion of Sections IV and V argues for this way of understanding the matter,

### Partially Visible Objects

I have been arguing that visual segregation is not sufficient for object-seeing. It is not necessary either.

Dretske thought of visual differentiation in terms of contour detection, but in fact it can take other forms. For instance, imagine a white paper marked with large blue polka dots. Now overlay on this a thin white paper with cut-outs that precisely match and are laid on top of the polka-dots. Looking at this display, you seem to see a single paper; it is just as in the *Camouflage* case, where you seem to see just one object, the naked bark of a tree. But now imagine that the overlay moves. The polka-dots do not stay covered; the motion of the overlay cuts off their contours and reveals some of the surrounding white. More strikingly, what you see is coordinated movement—all of the polka dots are cut-off simultaneously and in parallel. Because the dots are covered in unison, they look to be marks on a single rigid object distinct from the overlay. Thus, by seeing the movement, the overlay and polka-dot papers are visually differentiated from each other—you come to see two objects. Yet you do not see the contours of the lower paper; you do not segregate it from its surroundings. The differentiation is by relative motion, not by contour segregation. A lot of object-seeing is like this. In the *Camouflage* case, you don’t see the moth, but if it were to flutter its wings, you would be aware of seeing it for the moment that its wings moved. In these cases, causally coordinated movement indicates an object—we do not have to see its contours.

We also differentiate objects from their surroundings by perceptually learned capacities for discernment. In the famous low-resolution, high-contrast picture of a Dalmatian dog walking in a dappled wood, you at first see nothing but a random pattern of monochrome patches. But once your visual system deciphers the picture for what it is, you are consistently able to see the picture of the dog. Along the same lines, suppose that you are looking at a pond. You see no object in the water at first, but then you suddenly see the coordinated movement of a pattern of dots. Let’s suppose that through a process of perceptual learning and/or habituation, you have come to recognize this as the movement of a fish. Then, you might not merely see the movement of a two-dimensional pattern, but rather the three-dimensional undulation characteristic of a fish swimming. You may do this even though the outline of the fish is not visible.

### Non-Standard Visual Impressions

In addition to the straightforward cases in which we have an unobstructed, well-illuminated, middle-distance view of standard objects such as the above, we often see non-standard things in non-standard circumstances. It is not always clear whether this is object-seeing. When I look in a mirror, do I see myself as an object? When I look at a photograph of my grandmother, do I see *her* as an object? My proposal that straightforward object-seeing has a characteristic phenomenology helps us sort out these more perplexing cases—it doesn’t always give a straightforward yes or no answer, but it does help us figure out how these cases are different or functionally deficient.[[17]](#footnote-17)

*Mirror Images* Start with mirror-images. When you look at yourself in a mirror, you have a visual impression of an image of yourself located about two metres in front of you. That is,

(a) you are aware of a visually segregated region, *I*, of your visual field, and

(b) *I* is an image of you.

Let me say right away that it goes beyond my aims in this paper to offer an analysis of the claim (b), namely that *I* is an image of you—I’ll just rest with the notion that *I* projects a similarly coloured outline-shape as you would have, if (*per impossibile*)you had been standing two metres in front of yourself. (If this doesn’t fit your preferred notion of an image, substitute one that you prefer.) The important thing for my purposes is that *I* is a visually segregated part of the visual field. The question I want to ask is whether visual awareness of *I* amounts to object-seeing, specifically object-seeing yourself.

Now, the mirror-image *I* is not causally responsible for your visual state; it is not material and thus it is not causally responsible for anything. But since there is a relevant causal connection between your own body and what you see, many (including Grice and Roy Sorensen 2008) say that you see yourself in the mirror, but not the image. This uses CTP for the purpose to which it is best suited—to select among candidate-objects the one that is really the cause of your visual impression. But it ignores a different question. When you look in the mirror, do you *object*-seeyourself?

Yetter-Chappell (2018) observes that there is no essential difference between the causal process in which light travels directly from object to retina and one in which it bounces off a mirror.[[18]](#footnote-18) From this, she concludes that we “literally” see objects in mirrors. I don’t want to take issue with her premise, but this does not settle my question: do you object-see yourself in a mirror? That is: does the image *I* look as if it is a material object? I think not: it lacks some crucial characteristics of an object.

First, using a mirror depends on having a visual impression of a mirror over and above impressions of reflected objects. I have on occasion been startled in the murkiness of late evening by the proximity of someone, who, as it turned out, was me reflected in glass. Because I could not see the glass, what I saw seemed to be another person some distance in front of me.This suggests that there is a difference between cases where you can see the reflector and those in which you cannot: in the former case, you recognize the image *I* as a reflection in a mirror; in the second you do not. But this points to a significant difference between direct seeing and mirror seeing. In the latter case, you recognize *I* as a mirror image in part *by* seeing the mirror. In short, there is a manifest mediacy in the mirror-seeing case that is absent in the direct seeing case. Consequently, mirror-seeing is not object-seeing, or at least not direct object-seeing. Despite the similarity of the causal process leading up to my visual state, then, there is a difference in how we evaluate and understand the visual state.

Secondly, there is some evidence (by no means conclusive, in my estimation) that infants have trouble differentiating between shapes that are mirror-reversed—they tend to confuse p and q, b and d, etc., as well as left and right mirror-image profiles of the same face (Bornstein, Gross, and Wolf 1978, Dehaene et. al 2010). This suggests that they may be deficient with regard to performing the reversals necessary for gaining spatial information about objects from mirror-images. Consequently, they may not be able to use their eyes and their bodies properly to differentiate reflected objects from their surroundings—they would have trouble brushing a crumb off their own faces using a mirror. Again, animals view mirror images very differently: when these creatures are able to see the mirror, some (such as apes) see *some* object but don’t recognize it as themselves; some (such as, apparently, dogs) see no object at all. But presumably all of these animals are aware of *I*, a visually non-homogeneous region of their visual field.

Lastly, the perspectival dependencies of mirror images are different from those of real material objects. First of all, a mirror-reflection is tied to your own motion, but the closing velocity is twice your speed. This is a weird singularity: you can control and fine-tune your own actions by self-monitoring in a mirror (but only when you fully appreciate these perspectival changes and the left-right reversal).

These facts suggest that we *learn* to see in mirrors. We find it easy to brush our hair in the mirror, but we have learned to do this, just as some of us have learned to use microscopes, cameras in microsurgery, backup cameras in late-model cars, and so on. This does not suggest in the slightest that we do *not* see through these instruments. But it does suggest that we see things through them *indirectly*. It seems that we learn to see reflected objects by a special operation on a visually segregated part of the visual field.

In summary, there are two points that can be made here. The first is that whether or not you get information about yourself by looking at a mirror, there is something odd about the experience, though by now it’s so familiar that you don’t notice it. The larger point is that the *causal* similarities leading up to mirror seeing and real-world seeing are not decisive in deciding whether we see reflected objects as objects.

*Backlighting* Now let’s move to a more controversial case, Roy Sorensen’s (2008) treatment of solar eclipses. When the Moon totally eclipses the Sun, we see the Moon as a dark disc against the glow of the Sun’s corona. This much is agreed by all (including Sorensen himself). Sorensen, however, assumes that we can only see a three-dimensional material object such as the Moon by seeing its surface. And he asks: Which surface do we see? He says, provocatively, that we see its back surface—for this is the surface that is causally responsible for our visual impression. The *back* surface of the Moon is the one that intercepts the light from the Sun, thus giving rise to the impression of a dark disc.

Now, to talk of seeing surfaces is something of a departure from our discussion so far. We have been discussing object-seeing, not surface-seeing. But as a first approximation, let’s say that we surface-see *S* if (a) we are able to *differentiate* *S* from its surroundings by visible marks, and (b) we object-see the material object to which *S* belongs.[[19]](#footnote-19) By this criterion, we don’t see *any* surface of the eclipsing Moon, because no matter how hard we look, condition (a) is not met: there is simply no visual mark by which we can visually differentiate the surfaces of perfectly dark backlit objects from their surroundings. These surfaces are either in complete darkness (the facing surface) or hidden from us (the rear surface). We see an area of darkness, and, to state the obvious, we can’t see in the dark. However, in an eclipse, the silhouetted edge of the Moon is visible, because it coincides with inner limit of the lit-up surrounding area. So, it is possible to offer Sorensen an alternative way of maintaining his intuition (which, to be clear, I do not share) that material-object seeing is always indirect. For it seems appropriate to say that we see the Moon by seeing the edge of the Moon.

Taken in the spirit of Grice’s theory, then, Sorensen is correct: the rear surface of the Moon is causally responsible for a visual impression of a dark disc, and in the etiolated sense in which we see a perfectly camouflaged animal, we see it. But taken as a claim about object-seeing or surface-seeing, Sorensen is wrong. In Grice’s sense, what we see in an eclipse is an object, the Moon; we see it without seeing any part of its surface other than an edge (and we see the edge by seeing the illuminated surround). But we don’t object-see it.

*Depiction* Richard Wollheim (1973) argued that we see three-dimensional objects *in* pictures of them. I think he means something like this. Suppose that you are looking at a picture of object *O*,and:

(a) you are aware of a visually segregated region, *D*, of your visual field within the picture, and

(b) *D* depicts object *O*

Again, it is not part of my purpose to elucidate the depiction relation. Let’s just say that *D* depicts *O* if it is an image of *O*, i.e., if it projects a similarly-coloured outline shape as *O* would have.

Dominic Lopes (2005) gives an account of Wollheim’s “seeing in” that goes somewhat further than the above. He says that when we look at a picture, we have visual experience of marks, colours, and textures on a two-dimensional surface, and in virtue of this we have a visual impression as of a three-dimensional object. Kendall Walton (1984) takes this one step further by saying that since my long-dead grandmother is causally responsible for my visual impression of her through a photograph, I literally see her when I look at a photograph of her. Yetter-Chappell (2018) agrees, and she claims that the causal process in the production of photograph is sufficiently similar to that which is involved in direct seeing.

In my view, this reasoning goes wrong at an early stage—at least as it applies to object-seeing (and this appears to be Lopes’s intention). I do not have a visual impression as of a three-dimensional object when I look at a picture, because visual impressions are of objects as present to me here and now, and a picture of a three-dimensional object in no way makes it look as if the object is present (Hopkins 2012). The impression of *presence* is the feeling of a locus of physical interaction. (See Matthen 2005, chapter 13, for an account of the feeling of presence.) When I look at *D*,I have a visual impression of a two-dimensional pattern; when I examine it from different angles, my view of it changes in the manner of a two-dimensional object. Wollheim’s seeing-in is properly a matter of seeing a visually segregated region of the visual field, *D*,as a depiction of *O* (however depiction is to be analysed). I don’t see my grandmother as a material object in any picture of her, photographic or otherwise. All that I see in a picture is a depiction of her, a segregated portion of the visual field that projects the same similarly coloured outline shape as she would have done. So, the idea of literally seeing my grandmother doesn’t even get off the ground—if what is meant is that I object-see her.

Now, modify the case. What if we have a three-dimensional augmented reality presentation in which multiple films of my grandmother walking across a room are superimposed on the very room in which I standing by means of an augmented reality (AR) viewer that responds to my use of eyes and body—when I move around the room, the view changes accordingly, as does the view of my grandmother. It seems to me that in this case, by contrast with when I look at a two-dimensional picture or watch an ordinary three-dimensional movie, I genuinely have a visual impression as of a woman, and because of the causal connection that woman is my grandmother. In other words, I see *her* as an object in the AR presentation. But I don’t see her walking across the room: for me to see my grandmother walk across a room now, it would have to be the case that she is walking across the room now. Sadly, this cannot be so.

Like Sorensen, Walton might be right if the question is simply about the cause of a visual impression created when I look at the photograph. Perhaps, my grandmother was the cause of this visual impression. So, it might be that in the etiolated meaning attached to the term by sense-datum theory, I see my grandmother in a photograph of her. But I take it that Walton’s readers found this conclusion startling because they thought he was arguing that I see her in a more robust sense. And it is simply not the case that I object-see my grandmother (or any other three-dimensional object) when I look at a photograph.

### Conclusion

In this paper, I argued first that causal theories of perception are useful for deciding *which* object we see when we seem to see an object, but that they need to be supplemented by a notion of what it is to seem to see, and what it is to seem to see an object. I offered a phenomenologically accessible account of the first—to see is to have a veridical experience of things outside of oneself that is produced by looking. And I argued, without offering a specific account of its aetiology, that there is a characteristic phenomenology of seeing an object—it is to see a material object occupying a spatial region. My argument has offered evidence of visual phenomenology richer than that of sense-datum theory supplemented by visual segregation. Additionally, it offers common-sense purchase on ephemeral objects of vision such as mirror-images, backlit objects, and depictions.[[20]](#footnote-20)

**REFERENCES**

Aloiomanos, Yiannis, Weiss, Isaac, and Bandyopadhyay, Amit 1987. Active Vision. *International Journal of Computer Vision* 1 (4): 333-56

Austin, J. L. 1962. *Sense and Sensibilia* (reconstructed by G. J. Warnock). Oxford: Clarendon Press.

Balcetis, E., and Dunning, D. 2010. Wishful seeing: Desired objects are seen as closer. *Psychological Science* 21: 147-52.

Batty, Clare 2014. Olfactory objects. In D. Stokes, M. Matthen, & S. Biggs (eds.), *Perception and its Modalities*, Oxford: Oxford University Press: 222-46.

Bornstein, Marc H., Gross, Charles G., and Wolf, Joan Z (1978) Perceptual similarity of mirror images in infancy, *Cognition* 6: 89-116.

Brogaard, Berit 2015. Perceptual Reports. In Matthen (2015a): 237-53

Casati, Roberto 2015. Object Perception. In Matthen (2015a): 393-404.

Clark, Andy 2014. *Mindware: An Introduction to the Philosophy of Cognitive Science*. New York: Oxford University Press.

Chisholm, Roderick M. 1957. *Perceiving: A Philosophical Study* Ithaca N.Y.: Cornell University Press.

Churchland, Patricia S., Ramachandran, V. S., and Sejnowski, Terrence J. 1994. A Critique of Pure Vision. In C Koch and J. L. Davis (eds) *Large-Scale Neuronal Theories of the Brain.* Cambridge MA: Bradford Books, MIT Press: 23-62.

Clark, Austen 1996. Three Varieties of Visual Field. *Philosophical Psychology* 9 (4): 477-495.

Crowther, Thomas 2018. “In Touch With the Look of Solidity,” in T. Crowther and C. Mac Cumhaill (eds.) *Perceptual Ephemera* Oxford: Oxford University Press: 260-88.

Dehaene, S., Nakamura, K., Jobert, A., Kuroki, C., Ogawa, S. and Cohen, L. (2010) Why do children make mirror errors in reading? Neural correlates of mirror invariance in the visual word form area. *Neuroimage*, 49 (2):1837-48.

Dretske, Fred I 1969. *Seeing and Knowing* Chicago: University of Chicago Press.

Findlay, John M. and Gilchrist, Iain D. 2003. *Active Vision: The Psychology of Looking and Seeing* Oxford: Oxford University Press.

Fish, William 2010. *Philosophy of Perception: A Contemporary Introduction* New York: Routledge.

Fulkerson, Matthew 2014. *The First Sense: A Philosophical Study of Human Touch*. Cambridge MA: MIT Press.

Green, E.J. forthcoming. A Theory of Perceptual Objects. *Philosophy and Phenomenological Research* doi: 10.1111/phpr.12521

Grice, H.P. 1961. The Causal Theory of Perception I, *Proceedings of Aristotelian Society* supplementary volume 35: 121-52.

Grice, H.P. 1962. Some Remarks about the Senses. In R.J. Butler (ed) *Analytic Philosophy First Series*. Oxford: Basil Blackwell.

Hansen, T., M. Olkkonen, S. Walter and K.R. Gegenfurtner 2006. Memory modulates color appearance. *Nature Neuroscience* 9: 1367-68

Hopkins, Robert 2012. Seeing-In and Seeming to See. *Analysis* 72 (4): 650-59.

Jackson, Frank (1977) *Perception: A Representative Theory* Cambridge: Cambridge University Press.

Lewis, David 1966. Percepts and Color Mosaics in Visual Experience. *Philosophical Review* 75: 357-68.

Lewis, David 1980. Veridical Hallucination and Prosthetic Vision. *Australasian Journal of Philosophy* 58: 239-49.

Lopes, Dominic McIver 2005. *Sight and Sensibility.* Oxford: Oxford University Press.

Macpherson, Fiona 2012. Cognitive penetration of colour experience: Rethinking the issue in light of an indirect mechanism. *Philosophy and Phenomenological Research* 84 (1), 24-62.

Macpherson, Fiona. (ed.) (forthcoming). *Sensory Substitution and Augmentation,*Proceedings of the British Academy, Oxford University Press.

Martin, M.G.F. 2010. What’s In a Look? In B. Nanay (ed.)*Perceiving the World* New York: Oxford University Press: 160-225.

Matthen, Mohan 2005. *Seeing, Doing, and Knowing: A Philosophical Theory of Sense Perception.* Oxford: Clarendon Press.

Matthen, Mohan 2010. On the Diversity of Auditory Objects. *Review of Philosophy and Psychology* **1**: 63-89.

Matthen, Mohan 2014. How to Be Sure: Sensory Exploration and Empirical Certainty. *Philosophy and Phenomenological Research* 88 (1): 38-69.

Matthen, Mohan 2015a. *Oxford Handbook of the Philosophy of Perception* (ed.) Oxford: Oxford University Press.

Matthen, Mohan 2015b. Individuating the Senses. In Matthen (2015a): 567-86.

Maxwell, Grover 1962. The Ontological Status of Theoretical Entities. *Minnesota Studies in the Philosophy of Science* 3: 3-27.

Nudds, Matthew 2010. What are auditory objects? *Review of Philosophy and Psychology* 1(1): 105-122.

O’Callaghan, Casey 2007. *Sounds: A Philosophical Theory*. Oxford: Oxford University Press.

O’Callaghan, Casey 2016. Objects for multisensory perception. *Philosophical Studies* 173: 1269-89.

Price, H.H. 1932. *Perception.* London: Methuen.

Pylyshyn, Zenon 1989. The role of location indices in spatial perception: A sketch of the FINST spatial index model. *Cognition* 32 (1): 65-97.

Siegel, Susanna 2006. Subject and Object in the Content of Visual Experience. *Philosophical Review* 115: 355–88.

Siewert, Charles 1998. *The Significance of Consciousness.* Princeton: Princeton University Press.

Smith, Barry C. (2015) “The Chemical Senses,” In Matthen (2015a): 314-52.

Snowdon, Paul 1980-81. Perception, Vision, and Causation, *Proceedings of Aristotelian Society* NS 81: 175-92.

Sorensen, Roy 2008. *Seeing Dark Things* Oxford: Oxford University Press.

Stokes, Dustin 2012. Perceiving and Desiring: A New Look at the Cognitive Penetrability of Experience. *Philosophical Studies* 158 (3): 479-92.

White, Alan R. 1961. The Causal Theory of Perception II, *Proceedings of Aristotelian Society* supplementary volume 35: 153-68.

Yetter-Chappell, Helen (2018) Seeing through eyes, mirrors, shadows, and pictures. *Philosophical Studies* 175/8: 2017-42. DOI 10.1007/s11098-017-0948-8.

Walton, Kendall 1984. Transparent Pictures: On the Nature of Photographic Realism. *Critical Inquiry* 11: 246-77.

Wollheim, Richard 1973. The Work of Art as Object. In Wollheim, *On Art and the Mind: Essays and Lectures* (London: Allen Lane), pp. 112-29.

Xu, Fei 1997. From Lot's wife to a pillar of salt: Evidence that physical object is a sortal concept. *Mind & Language* 12 (3‐4): 365-92.

1. Thomas Crowther (2018) discusses a closely related topic, the visual appearance of solidity. [↑](#footnote-ref-1)
2. Many authors (for example, Fish 2010) take the notion of a visual impression for granted and are happy simply to stipulate that CTP applies only to these. There is nothing wrong with this; elucidating vision is not the purpose of CTP. Nevertheless, an adequate elucidation of visual impressions, and particularly of object-impressions, provides a fuller understanding of the phenomena mentioned at the beginning of this paper and helps scotch some wildly counter-intuitive applications of the theory. [↑](#footnote-ref-2)
3. Despite his title, “The Causal Theory of Perception,” Grice was not primarily interested in perception. His aim, and his great achievement in this paper, was to distinguish between semantical and pragmatic explanations of everyday language implications. The main example, made relevant by the then unpublished views of J.L. Austin (1962) and of others listed by Alan White (1961, 158), was the idea that ‘*X* looks red’ implies either that somebody has denied that it is red, or that there is cause for doubt that it is. Grice transformed philosophy of language with his argument that this is a pragmatic, not a semantic, implication. [↑](#footnote-ref-3)
4. Paul Snowdon (1980-81) says that ‘looks’ should be understood “phenomenologically” in Grice’s CTP. I assume he is alluding to Frank Jackson (1977), who defines this as ‘looks *F*’ where *F* is a term for a colour, shape, or distance. This overlaps with the proposal considered in section III, and I’ll discuss it there. [↑](#footnote-ref-4)
5. To conform to Grice’s formulation, I have substituted ‘looks a certain way’ for ‘sees’ in this definition. As well, I have not (as Dretske does) insisted on *logical* necessity, which cannot, by its nature, apply to some but not all statements of the same logical form, and cannot therefore distinguish between ‘My hand looks this way’ and ‘His finances look this way.’ I have substituted the vaguer term, ‘analytically.’ [↑](#footnote-ref-5)
6. There are other accounts of the semantics of ‘looks’ as it is used to report visual impressions. See, in particular, Martin (2010). For a sophisticated overview, see Brogaard (2015). [↑](#footnote-ref-6)
7. Such looks are a problem for artificial intelligence recognition-modules: the reason why it takes advanced computational techniques to recognize faces is that the look of a face is *not* reducible to low-level visual properties. (See Martin 2010 for an explication of *looks*.) [↑](#footnote-ref-7)
8. Snowdon (1980-81, 176) incorporates visual phenomenology in his formulation of CTP, but he retains the tie to reports of the form, “It looks to me as if . . .” This reluctance to talk directly of visual experience is now *passé*: see, for example, Fish (2010), chapter 7. But it needs to be clarified that not every experience characteristic of seeing counts in this context: being dazzled by a flash bulb and having an after-image of it are effects of seeing the light, but do not constitute seeing it. For this reason, I have specified that it is an experience as of things outside the perceiving subject. [↑](#footnote-ref-8)
9. The same goes for the expectations that go with active looking. Suppose that the content of seeing something as a 3D object includes the expectation that if you move to the left, you will see parts of it that were theretofore hidden from sight; suppose the content of seeing something as a shadow or stain similarly includes the expectation that you will not. Then, the content of *imagining* that you see such objects includes the same expectations. [↑](#footnote-ref-9)
10. I should say here, for the sake of clarity, that attentive looking *at an object* is not necessary in order to see *that object*. You may look at a scene, or at one object in a scene, and thereby see (other) objects that you are not directly looking at. [↑](#footnote-ref-10)
11. For discussions of active vision, see Aloiomanos et. al (1987), Churchland et. al. (1994), Findlay and Gilchrist (2003), Clark (2014), 101; and Matthen (2014). [↑](#footnote-ref-11)
12. Sensory substitution is an interesting case. TVSS-stimulated blind people who are tactually stimulated do report visual-like experiences of depth and perspective produced by an activity that resembles looking in relevant ways. I won’t pursue this discussion here, but see Macpherson (forthcoming), especially the Introduction. [↑](#footnote-ref-12)
13. Embarrassingly for Grice, he would have to say that the colour-blind person does see the same ‘5’ as the colour-sighted person. [↑](#footnote-ref-13)
14. Grice writes, “If someone has seen a speck on the horizon, which is in fact a battleship, we should in some contexts be willing to say that he has seen a battleship” (147). True: a lookout would get credit for having spotted it. But ist’s hard to extend the same courtesy to a spotter who failed to discern a camouflaged moth right before his eyes. [↑](#footnote-ref-14)
15. Many thanks to Maarten Steenhagen who showed me some astonishing drawings of migraine auras, showing striking similarities between different subjects. [↑](#footnote-ref-15)
16. It’s not clear to me how his account would work in olfaction, gustation, and touch. [↑](#footnote-ref-16)
17. In a clever and instructive paper, Helen Yetter-Chappell (forthcoming) explicitly reverses this methodology, arguing (see her note 3) for an approach in which examples drive theory, rather than the other way around. I agree with many of the conclusions that she arrives at, but my aim to show how things go wrong or are different in certain examples, and here the theory of object-seeing helps. [↑](#footnote-ref-17)
18. There is a distant ancestor of this argument in Maxwell (1962), who asks whether we see “corporeal organisms” through microscopes, observing that there is a continuum (of unspecified dimensionality) between seeing through a vacuum and seeing through a microscope. It’s unclear, though, why either Yetter-Chappell or Maxwell before her, take the similarity of the *external* process to be the sole determinant of whether or not we see an object. Why don’t they allow that the dissimilarity of the resultant visual impression is relevant? [↑](#footnote-ref-18)
19. If we did not see the material object to which *S* belongs, there would be no difference between visually segregating *S* and seeing it as a material surface. [↑](#footnote-ref-19)
20. I was privileged to present versions of this paper at Oberlin College, New York University (Abu Dhabi), and Durham University. I am very grateful to Todd Ganson, Gabe Rabin, and Clare Mac Cumhaill for these opportunities and for extremely helpful discussion. I thank my audiences for discussion, particularly my commentator in Abu Dhabi, Phillip Meadows. [↑](#footnote-ref-20)