
MENTAL IMAGERY AND THE VARIETIES OF AMODAL PERCEPTION

BY

ROBERT EAMON BRISCOE

Abstract: The problem of *amodal perception* is the problem of how we represent features of perceived objects that are occluded or otherwise hidden from us. Bence Nanay (2010) has recently proposed that we amodally perceive an object's occluded features by imaginatively projecting them into the relevant regions of visual egocentric space. In this paper, I argue that amodal perception is not a single, unitary capacity. Drawing appropriate distinctions reveals amodal perception to be characterized not only by mental imagery, as Nanay suggests, but also by genuinely visual representations as well as beliefs. I conclude with some brief remarks on the role of object-directed bodily action in conferring a sense of unseen presence on an object's occluded features.

1. Introduction: interactions between imagination and perception

Seeing and imagining engage overlapping neural machinery and interact in a variety of important ways. In particular, there is evidence that forming mental images evokes retinotopically organized activation patterns in early visual processing areas (Farah *et al.*, 1992; Kosslyn *et al.*, 1999; Slotnick *et al.*, 2005) and that visualizing an object can facilitate its subsequent recognition (Kosslyn and Sussman, 1995).¹

Human beings and perhaps other animals also have the ability to 'augment' reality by superimposing mental imagery on the visually perceived scene.² Thus, as Rick Grush observes, one might use a mental image to decide where in egocentric space a vase should be placed in order best to obscure a picture on one's desk (2004, p. 390). Similarly, when deciding how to arrange the furniture in a new home, one might imaginatively project an armchair into an empty corner or a painting onto an unadorned wall. Yet another familiar application of this ability is the experience of

noticing a constellation in the nighttime sky. Noticing a constellation is a hybrid, visual-imaginative experience: it involves both *seeing* the stars in the constellation and *imagining* the lines that connect them at the same time. In what follows, I shall refer to such hybrid experiences – involving both a bottom-up, externally generated component and a top-down, internally generated component – as ‘make-perceive.’³

Bence Nanay (2010) has recently suggested that our ability to project mental imagery finds an important application in our everyday visual experience of opaque, 3-D objects. According to Nanay, ‘the exercise of mental imagery is necessary for amodal perception: for representing those parts of perceived objects that are not visible’ (p. 240). Thus, when I see a cat whose tail is hidden behind a fence picket, I may imaginatively project a mental image of its occluded tail on a certain location in my visual egocentric space. Since Nanay, like other recent philosophers, including Alva Noë (2004, 2005) and Sean Kelly (2004), assumes that seeing 3-D objects almost always involves amodal perception, imagination turns out to be, on his view, a ‘necessary ingredient’ of perception itself.⁴

In developing this Kantian-Sellarsian proposal,⁵ Nanay rejects two other accounts of how we represent non-visible object-features: a perceptual account, according to which we in some sense see them, and a belief-based account, according to which we do not see them, but rather infer their unseen existence. In defending his imagery-based account, Nanay also poses objections to Alva Noë’s enactive, sensorimotor account of amodal perception, according to which we do not in fact *represent* features of objects that are not currently visible to us, but rather deploy practical knowledge of how to bring those unperceived features into view by moving our bodies. It is our possession of such practical, sensorimotor knowledge, Noë says, ‘that is the ground of the claim that, in seeing . . . the visible parts of the cat, you have a *visual* sense of the presence of the strictly unseen, or strictly *invisible* parts of the cat’ (2005, pp. 242–43).

I agree with Nanay that Noë’s account of amodal perception faces a number of serious problems. That said, I think that Nanay and Noë alike are mistaken in assuming that amodal perception can be profitably approached as a single, unitary capacity. Amodal perception, I shall argue in this paper, is not a psychological natural kind. Rather, it includes both cases of stimulus-driven or ‘non-cognitive,’ amodal completion (*NA-completion*) as well as cases of top-down or ‘cognitive’ completion (*C-completion*). The latter form of completion involves very different representational resources and processing mechanisms than the former, and gives rise not only to mental imagery, but also to beliefs. In Section 2, I begin by characterizing some main differences between NA-completion and C-completion. In Section 3, I argue that NA-completion is a properly *perceptual* phenomenon and that NA-completed contours and surfaces are not usefully characterized as ‘lower-level’ mental imagery. According to

Nanay, the belief-based account does not provide even a partial explanation of amodal perception. In Section 4, I show that Nanay's argument for this conclusion is unconvincing. There are plausibly many common cases in which we form beliefs about a perceived object's non-visible features on the basis of its currently visible features and/or background knowledge. I conclude in Section 5 with some brief remarks on the role played by object-directed *bodily actions* in conferring a sense of unseen presence on the hidden features of the objects that we perceive.

2. *Representing the occluded features of perceived objects*

Nanay defines 'amodal perception' as the representation of those parts of perceived objects that are not currently visible. (This may be because they are, e.g., occluded, or not illuminated, or project onto the retinal blind spot.) Like Noë (2004, 2005), Nanay assumes that amodal perception occurs whenever we see an opaque, 3-D object: 'Almost all episodes of perception include an amodal component. For example, typically, only three sides of a non-transparent cube are visible. The other three are not visible – we represent them "amodally." The same goes for houses or for any ordinary object. We perceive the back side of any (nontransparent) object only amodally' (p. 242).

Like Noë, Nanay also subsumes long-studied *amodal completion* phenomena (Michotte *et al.*, 1964/1991) under the rubric of amodal perception.⁶ A standard way of drawing distinction between 'modal' and 'amodal' completion by students of perception is as follows. In modal completion, the observer characteristically has a distinct, quasi-visual impression of a contour or surface where there are no corresponding stimulus features in the retinal image. Natural scenes that most commonly give rise to modal completion are those in which a foreground surface is *camouflaged* by a more distant background surface. 'Because the boundaries of the camouflaged object do not project any contrast,' Fleming and Anderson (2004) write, 'they have no corresponding features in the image and thus the nearer object is effectively invisible. Under these circumstances, the visual system must actively "hallucinate" the invisible structures' (p. 1288). As a familiar example of this type of completion, consider the illusory Kanizsa squares in Figure 1a. Most observers, when viewing the figure, report having the impression of seeing a 'thin' square on the left (in which the illusory, vertical contours bow inwards) and the impression of seeing a 'fat' square on the right (in which the illusory, vertical contours bow outwards). The interpolated illusory squares exemplify the 'phenomenal filling-in' (Pessoa *et al.*, 1998) characteristic of modal completion.

By contrast, 'amodal' completion occurs when one object appears to be partially occluded by another and does not typically result in a

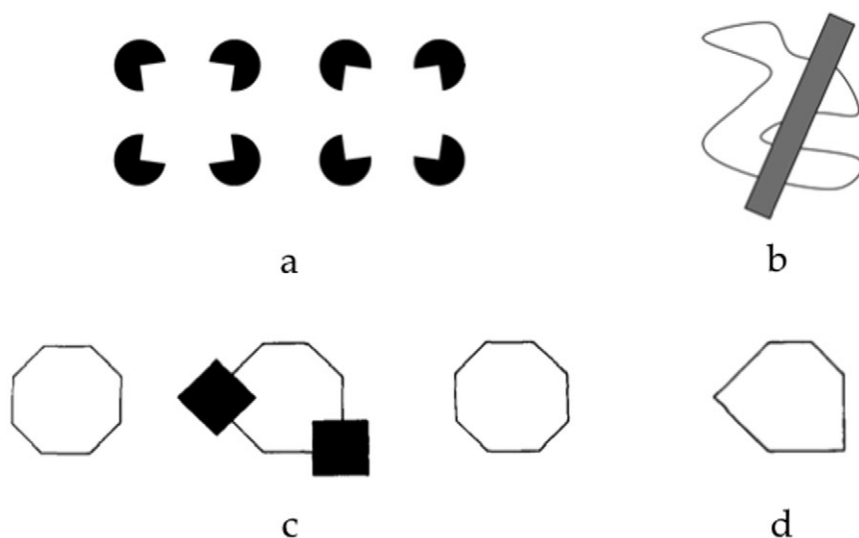


Figure 1 Modal completion and amodal completion.

quasi-visual impression of the object's hidden contours or surfaces. Amodal completion is not characterized by phenomenal filling-in. (Hence, the epithet 'amodal.') Rather, the phenomenally most salient characteristic of amodal completion is the perceived *unity* of the partially occluded object (Michotte, 1964/1991; Kanizsa, 1979). Thus, when viewing Figure 1b, we see what appears to be a single, white surface partially occluded by a grey strip. That is, a grey strip appears to be closer in depth and to hide parts of a white surface that complete behind it. Our visual impression is not of four, unrelated image regions on the same plane of depth.

As another example of amodal completion, consider Figure 1c. Notice here that the visual system completes the occluded object in the manner indicated in Figure 1d, even though the context provided by the flankers would rather suggest that the object is an octagon. This example nicely illustrates the point that the interpolation process in amodal completion 'follows complex principles of its own' (Pylyshyn, 1999, p. 345) and is not rationally sensitive to the observer's beliefs and other high-level cognitive states (Kanizsa, 1985; Kamitani and Shimojo, 2004). Similar points, I should note, also hold with respect to modal completion. Like the shapes of amodally completed contours, the shapes of modally completed contours are the result of stimulus-driven operations that do not depend on stored, object-specific knowledge and are not subject to top-down control.⁷ (Thus, in Figure 1a, rotating the pac-man-shaped inducers slightly

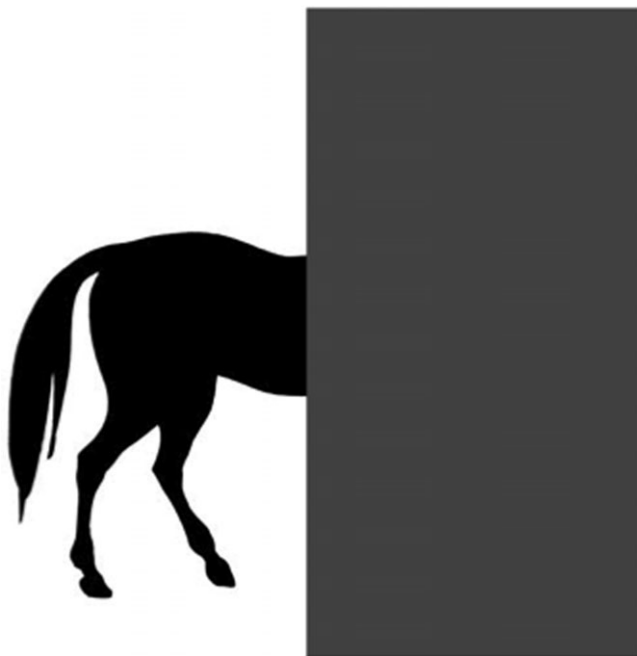


Figure 2 Cognitive completion.

inwards automatically transforms the ‘thin’ square on the left into the ‘fat’ square on the right. An observer cannot *will* herself to see illusory contours with different shapes than these.)

The main point, for present purposes, is that amodal and modal completion are stimulus-driven processes subserved by relatively low-level, vision-specific, neural machinery. Such ‘non-cognitive’ characterization of amodal and modal completion (Kanizsa and Gerbino, 1982), is supported by empirical evidence that amodally and modally completed contours are represented by bottom-up cell-activations in early visual processing areas such as primary visual cortex (Sugita, 1999; Bakin *et al.*, 2000; von der Heydt, 2004). In what follows, I shall accordingly refer to the kind of completion exhibited in Figures 1b and 1c as non-cognitive, amodal completion (or *NA-completion* for short) and to the kind of completion exhibited in Figure 1a as non-cognitive, modal completion (or *NM-completion* for short).

Now consider Figure 2. When an observer represents the hidden parts of the occluded animal in the figure (a horse facing to the right), the features that she attributes to them are likely to be constrained by her past experiences and beliefs. It is unlikely, e.g., that the observer will represent the rectangular occluder as concealing the front half of an armadillo, or an

alligator, or a speedboat. The kind of completion illustrated by Figure 2 and by the cat example above, in short, has an intuitively cognitive or 'top-down' component that is absent in the kind of completion illustrated by Figures 1a–1c. In particular, it characteristically involves *identification* of the perceived object and, hence, accessing categorial information stored in long-term memory. (Hence, to use the terminology of Machery [2009], it is a 'higher cognitive competence'.)

In what follows, I shall refer to the kind of completion exhibited by Figure 2 as cognitive completion (or *C-completion* for short). C-completion is characterized here as 'cognitive' not because it is assumed to result in beliefs or other non-perceptual representations (though it may well), but rather because it is assumed to *depend* on the observer's beliefs and background knowledge. This is to say that C-completion processes are cognitively penetrable: unlike NA- and NM-completion processes, they can be 'altered in a way that bears some logical relation to what the person knows' (Pylyshyn, 1999, p. 343).⁸

According to Nanay, when we represent features of perceived objects that are not currently visible to us, we 'use mental imagery . . . in a way that would allow us to localize the imagined object in our egocentric space' (p. 250), i.e. we engage in what I above called 'make-perceive.' What kind of completion is involved when we make-perceive an object's non-visible features, e.g., a cat's occluded tail? To the extent that processes involved in the formation of the projected mental image are guided by information about the object stored in long-term memory, it clearly seems to involve C-completion. This assessment is supported by Nanay's acknowledgment that 'amodal perception relies heavily on our background knowledge of how the occluded parts of the object (may) look. If I have never seen a cat, I will have difficulties attributing properties to its tail behind the fence' (2010, p. 247).⁹ Amodal perception, however, only relies on observers' background knowledge in cases of C-completion. Unlike NA-completion, C-completion is not a stimulus-driven process. It also depends on stored information about the kind of object we are viewing and/or its specific features as an individual.

As mentioned above, Nanay rejects two alternative accounts of the way in which we represent the hidden features of the objects that we perceive. The first is that we actually perceive them (the perceptual account). The second is that we do not see them, but rather infer their existence from visible object-features (the belief-based account). In what follows, I argue that distinguishing between NA- and C-completion reveals that both perception and belief contribute to the representation of non-visible object features. I first show, in Section 3, that NA-completion mechanisms produce properly *perceptual* representations of occluded surfaces in early visual processing areas. I then show, in Section 4, that C-completion mechanisms produce not only mental imagery, but also *beliefs*. Contrary

to Nanay, the imagery-based account provides only a partial explanation of amodal perception.

3. *NA-completion is perceptual*

Nanay attributes a perceptual account of amodal perception to J. J. Gibson. He quotes Gibson as saying that ‘the perception of occlusion . . . entails the perception of *something* which is *occluded*’ (1972, p. 229). Nanay’s objection is that an observer cannot be literally said to see an object or object-feature that does not optically project onto the retinal surfaces in her eyes. ‘[A] relatively widely accepted necessary condition for perception,’ he writes, ‘is the presence of sensory stimulation’ (p. 242). Since occluded surfaces are, as Gibson would put it, optically ‘unprojected,’ the perceptual account seems obviously false, if not self-contradictory.

Can the perceptual account be so easily dismissed? I agree with Nanay on independent grounds that we do not *see* C-completed object-features, like the cat’s hidden tail or the front half of the horse in Figure 2. I also agree with Nanay that we do not consciously see NA-completed contours and surfaces. These are described as ‘amodal’ for good reason. That said, I think that an empirically and phenomenologically compelling case can be made for the view that NA-completion, unlike C-completion, is a properly *perceptual* phenomenon subserved by representations of occluded object-features in early visual processing areas.

Gibson did more than any other vision scientist of the twentieth century to highlight the importance of occlusion for how we perceive the way in which environing surfaces are arrayed around us in depth (Gibson, 1966, 1979). Nonetheless, Gibson’s account of what we see when one object partially conceals another object or when the front side of an object conceals its backside may appear somewhat paradoxical. One the one hand, he allows that occluded surfaces are optically unprojected and, in this sense, ‘out of sight or hidden from view’ (1979, p. 78). On the other hand, he claims that occluded surfaces may be perceived under certain circumstances. Is there a genuine paradox here?

The appearance of paradox, I would suggest, can be dispelled if the *informational basis* for the perception of a surface is not limited to the surface’s optical projection in the retinal image (and the sensory stimulation caused thereby). And this is just what Gibson centrally maintains: there are a number of independently variable sources of visual information for occlusion, information that ‘specifies the existence of one surface behind another, i.e. the continued existence of a hidden surface’ (1966, p. 204). In what follows, I shall argue that, the visual system uses

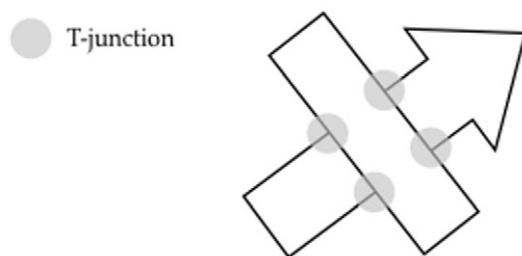


Figure 3 T-junctions.

information for occlusion in order to form representations of occluded object-features that contribute both to the phenomenology and content of conscious visual experiences.¹⁰

One important source of information for occlusion comes from the deletion or ‘wiping’ of surface texture under perspective transformation:

Koffka asserted . . . that he could ‘see’ the top of his table extending behind the book that lay on it. He considered this an example of the figure-ground phenomenon. . . . But insofar as his head moved and the texture of the table was wiped and unwiped by the edges of the book, he had information for perceiving the table behind the book (Gibson, 1966, p. 204).

Our visual experience of one thing as behind another in this situation, Gibson says, ‘is not a paradox, for there is information to specify it’ (1966, p. 204). Other sources of information for occlusion studied by contemporary vision scientists include binocular disparities that arise due to the lateral separation of the eyes (stereopsis) and ‘T-junctions,’ formed where an occluding edge intersects an occluded edge in the visual image (Figure 3).¹¹ Binocular disparities and T-junctions both play an important role in 3-D surface representation: they enable the visual system to estimate which of two surfaces that meet at an optically specified edge is closer in depth and therefore ‘owns’ the border between them. Since surfaces that do not own borders are effectively unbounded, they can connect to other unbounded surfaces and amodally complete behind their occluders (Nakayama *et al.*, 1995; Fleming and Anderson, 2004).

There is a growing body of data that early visual processing areas such as V1 (primary visual cortex) and V2 use information for occlusion for purposes of NA-completion (Sugita, 1999; Bakin *et al.*, 2000; von der Heydt, 2004). Sugita (1999), e.g., found that NA-completion in V1 is modulated by binocular disparity. Orientation-selective cells in V1 were presented with two vertical line segments separated by a gray patch. When the patch was presented with zero disparity or uncrossed disparity, so that it appeared, respectively, on the same or a more distant plane of depth than the line segments, the cells did not respond. However, when the patch

was presented with crossed disparity, so that it appeared to be in front of the line segments – a stimulus consistent with *occlusion* of a single, vertical bar – the cells responded vigorously. The notion that the visual system exploits information for occlusion for purposes of constructing perceptual representations of occluded object-features thus seems to be on firm empirical footing.¹²

It should be emphasized that although representations of occluded object-features formed in early visual processing areas do not result in the ‘filling-in’ characteristic of NM-completion, they nonetheless make an important contribution both to the phenomenology and content of conscious, visual experience. In particular, they play a significant role in the *organization* of the 3-D visual scene that we perceive. This point can be brought out by reflection on Figure 4a. When we look at the figure, our visual impression is of eight, separate, 2-D image fragments on the same plane of depth. We may with prolonged inspection imagine the fragments as joining to form a Necker cube, but that is not what we in the first instance see. Next consider Figure 4b. Our visual impression of the presence of partially occluded Necker cube is now quite vivid. I think it is fair to say that we see the visible fragments of the cube as *connected* – that we see the cube in the figure as a single, discrete object – even though we receive no sensory stimulation from the regions of the cube intervening between those fragments. This example nicely illustrates the point that the information available to the visual system for the existence of an object-feature is not limited to the feature’s projection in the retinal image. It also illustrates the point that representations of occluded object-features formed on the basis of information for occlusion play an important role in the consciously experienced, perceptual organization of the 3-D scene.

Another instructive example is provided by Figure 5. Figure 5a is bistable: with prolonged viewing the apparent depth relationships can *flip*, causing the units in the figure to alternate between a modal and an amodal appearance.¹³ This alternation is illustrated in Figures 5b and 5c. In Figure 5b, an illusory diamond modally completes in front of four amodally completed black disks. In Figure 5c, by contrast, the diamond amodally completes behind four modally completed holes. As Kellman and Shipley (1991) point out, the shapes of some of the modally and amodally interpolated boundaries in Figure 5b and Figure 5c are the same. There are some important differences, however, in the overall perceptual organization of the 3-D scene (Anderson, *et al.*, 2002). In Figure 5b, we have a visual experience as of a modally completed, white square; four, amodally completed, black discs; and a white surface in the background. By contrast, in Figure 5c, we have a visual experience as of a white surface in foreground; four, modally completed, circular holes; an amodally completed, white square; and a black surface in the background. It seems clear

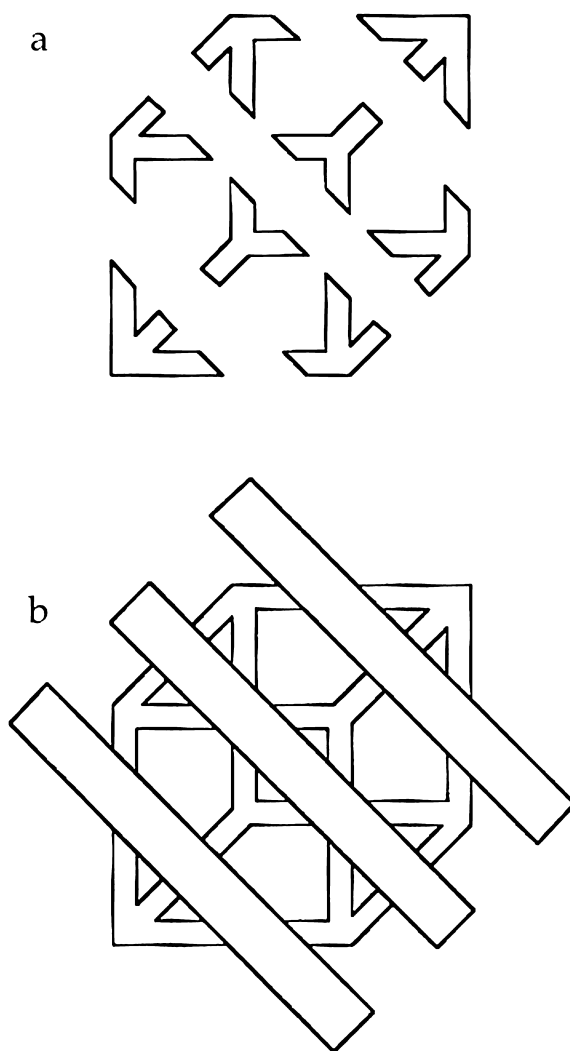


Figure 4 Completion and object segmentation.

that two phenomenologically different scenes, containing different objects, are perceived.

The last pair of figures illustrate the point that NA-completed contours and surfaces are not non-perceptual addenda to what we ‘strictly speaking’ see. Rather, NA-completion plays an integral role in perceptually segmenting the visual scene into discrete, 3-D objects at different distances in depth (Nakayama *et al.*, 1995; von der Heydt, 2004; Fleming and Anderson, 2004). Indeed, without the stimulus-driven processes that result in

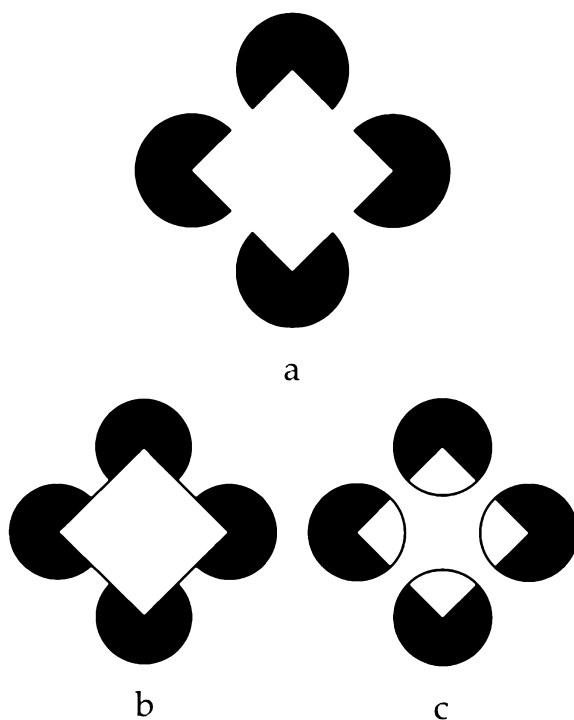


Figure 5 Completion and visual scene organization. (Adapted with permission from Anderson, *et al.*, 2002.)

NA-completion (and NM-completion), the question of how we represent the occluded features of discrete, 3-D objects would not arise because we would not *see* coherently organized scenes consisting of such objects.

According to Nanay, we use mental imagery in order to represent the occluded features of the objects that we perceive. This has the implication that we do not perceive the *unity* of the partially occluded cube in Figure 4b, as I have here suggested, but rather only imagine it. Imagination is the ‘mental glue,’ as it were, that binds the visible parts of partially occluded objects together into discrete, coherent wholes. We have now obtained three important conclusions, however, that militate against this view:

First, the informational basis for the perception of a contour or surface is not limited to its optical projection in the retinal image. Texture deletion under perspective transformation, binocular disparity, and T-junctions are all sources of ecologically valid visual information for the existence of occluded object-features. The absence of sensory stimulation from a feature *F* thus does not always mean the absence of visual information for *F*. Second, there is empirical evidence that sources of visual information

for occlusion are actually used by NA-completion mechanisms to form representations of occluded object-regions in early visual processing areas such as primary visual cortex. Last, the representational outputs of NA-completion mechanisms play a fundamental role in the perceptual grouping of visible surface regions in depth. In particular, they help causally to explain our visual experience of partially occluded objects as coherent, unified wholes.

Nanay is not alone in suggesting that we represent the occluded features of perceived objects in NA-completion using mental images. Lehar (2003), e.g., suggests that 'amodal perception is actually a primitive, or lower-level form of mental imagery' (p. 187). It may thus be useful at this juncture to review some important differences between mental imagery and NA-completion phenomena:

Nanay writes that 'if what it is like to have visual imagery is similar to what it is like to perceive and being aware of occluded parts of perceived objects is having visual imagery, then, putting these two claims together, we get that what it is like to be aware of the occluded parts of perceived objects is similar to what it is like to perceive those parts that are not occluded' (p. 252). NA-completion, however, is described as 'amodal' for good reason. In contrast with interpolated contours and surfaces in NM-completion (Figure 1a), interpolated contours and surfaces in NA-completion do not have a visual or quasi-visual phenomenology. What it is like to be aware of the occluded parts of Figure 1b, for example, is *not* similar to what it is like to be aware of the parts of the figure that are not occluded. The first difference, then, is that visual mental images, by hypothesis, have a conscious, visual or quasi-visual phenomenology, while NA-completion phenomena do not.

Second, mental images are not *stable* in the absence of sustained effort and fade rapidly. As Hume puts it, 'in the imagination the perception is faint and languid, and cannot without difficulty be preserv'd by the mind steady and uniform for any considerable time' (1739/2000, p. 11). By contrast, NA-completion phenomena normally persist so long as one perceives their inducers and are not subject to volitional control. No more effort is required to see the partially occluded Necker cube in Figure 4b than to see its occluders.

Third, mental image formation, unlike NA-completion, is not *stimulus-driven*. According to the currently most influential account of depictive mental imagery, mental images are formed when high-level object recognition systems (in inferotemporal cortex) project stored object-information backward into early visual processing areas (Kosslyn, 1994; Kosslyn *et al.*, 2005). By contrast, NA-completion does not require a high-level, top-down component, and utilizes many of the same low-level mechanisms that are involved in the perception and organization of non-occluded object-features.

Last, mental images are not *obligatory*. When I see a cat partially occluded by a picket fence, I may imagine the shape of its hidden tail, but I may also imagine any of variety of other things instead, or indeed nothing at all. By contrast, my experience of the NA-completed contours in Figures 1b, 1c, 4b, 5b, and 5c is not similarly subject to volition or top-down influence. NA-completion, far from exhibiting what Hume called the ‘liberty of the imagination’ (1739/2000, p. 12), operates automatically in accordance with a fairly strict set of organizational principles. Our experience of seeing distinct visible surfaces as belonging to a single, unified object in NA-completion, in consequence, is not pliant in the way that one would expect it to be were NA-completed contours and surfaces represented using top-down, mental imagery.

Given these four main differences, NA-completion phenomena, we can conclude, are not usefully characterized as a ‘lower level’ form of mental imagery.

4. *The belief-based account*

According to the belief-based account, a perceived object’s hidden features are represented by means of beliefs inferred from the object’s visible features as well as relevant background knowledge. One of Nanay’s objections to the belief-based account is that when the visual system NA-completes a partially occluded object, it does so in a way that is generally insensitive to our beliefs and expectations. I may believe, e.g., that the occluders in Figure 1c conceal an octagon, but the completion my visual system prefers is the asymmetrical shape shown in Figure 1d. This strongly suggests that the belief-based account of NA-completion is mistaken. Because cases of NA-completion are, by hypothesis, cases of amodal perception, it follows that the belief-based account fails to provide a general explanation of amodal perception. Call this objection the argument from the belief-insensitivity of NA-completion.¹⁴

The natural question to ask at this point is whether the belief-based account adequately explains certain cases of C-completion. Since C-completion processes *are* sensitive to background knowledge (indeed this is what makes them ‘cognitive’), and since C-completion is, by hypothesis, a form of amodal perception, it seems reasonable to suppose that the belief-based account provides a partial explanation of amodal perception.

As Nanay acknowledges, the argument from belief-insensitivity only shows that ‘there are *some* cases of amodal perception when the occluded parts of the perceived object are not represented by a belief’ (p. 244, my emphasis), namely, cases of NA-completion. In order to show that the belief account does not provide even a partial explanation of amodal perception, Nanay appeals to empirical evidence that NM-completed

contours are represented by cell activations in primary visual cortex (area V1) (Lee and Nguyen, 2001). Assuming (1) that the same neural mechanisms in early vision account for both NM- and NA-completed contours¹⁵ and (2) that the belief-account predicts that NM-completed contours are not represented in V1 ('If these shapes [those of NM-completed contours] are represented by a belief, then the primary visual cortex is not supposed to represent them' (p. 245)), it follows that 'the belief-view is unlikely to be the right way of explaining the representation of occluded shapes' (p. 246), i.e. NA-completion.¹⁶

The inadequacy of the belief-based account to explain representation of occluded features in NA-completion, however, was already established by the argument from belief-insensitivity above. The question was whether the belief-based account could adequately explain representation of occluded features in cases of *C-completion*. Empirical evidence that NA-completion has neural correlates in early visual processing areas seems consistent with answering this question affirmatively.

Are there reasons to think that the belief-based account *does* explain cases of C-completion? In order to show that there are, I take it that it is enough to show that there are cases in which we form beliefs about a perceived object's non-visible features on the basis of its visible features and/or background beliefs. And, indeed, relevant cases seem far from uncommon. For example, I may form the belief that there is rust damage somewhere on the hidden side of my neighbor's car on the basis of the rust damage visible on its facing surfaces and the assumption that the effects of weathering are unlikely to be confined to them. Similarly, if I believe that a painting on the wall in front of me was signed by the artist, but do not see a signature anywhere on the visible side of the canvas, then I may form the belief that there is a signature somewhere on its opposite side. No doubt, there are situations in which we first project a mental image of a perceived object's occluded features, in the manner suggested by Nanay, and then form beliefs about those features on the basis of what we imagine. But plausibly there are also many situations in which we form beliefs about a perceived object's occluded features *directly* on the basis of the object's visible features and collateral information, i.e. without first projecting a mental image of its occluded features. (Nanay adduces no evidence that this is psychologically impossible or even psychologically anomalous.) If this is right, however, then, contrary to Nanay, there are reasons to think that the belief-account does provide a partial explanation of amodal perception.

At this point, we have ample reason to reject the thesis that mental imagery is *necessary* for amodal perception. That said, nothing I say here should be taken to conflict with the assumption that projecting mental imagery in the manner suggested by Nanay is sufficient for amodal perception. If amodal perception is defined as representing a perceived

object's non-visible features, and if mental images are representational, then forming mental images of a perceived object's occluded features suffices for amodal perception. This modest thesis is empirically well motivated. In particular, there is neuropsychological evidence that feedback connections may enable high-level visual areas in the brain sometimes to 'augment' degraded perceptual inputs or, in the case of partial occlusion (superposition), incomplete perceptual inputs with stored, object-specific information (Kosslyn, 1994, ch. 6). Of direct relevance is Kosslyn's suggestion that top-down augmentations may be experienced as conscious, visual mental images coded in viewer-centered or 'egocentric' coordinates (Kosslyn, 1994, ch. 5). There is thus good empirical motivation for the claim that we sometimes represent occluded object-features by superimposing mental images on the relevant region of the visually experienced scene.

5. *Bodily action and motoric completion*

I have argued that the imagery-based account only explains some cases of C-completion and that it, in general, fails to explain NA-completion. I have also argued that amodal perception is not a single, unitary capacity – a natural, psychological kind as it were – but rather a motley, comprising perceptual representations in the case of NA-completion as well as beliefs and mental imagery in the case of C-completion.¹⁷ In this final section, I would like to suggest that representations of object-directed bodily actions also play an important role in amodal perception.

When we engage in visuomotor actions targeted on an object, the kinematic parameters of the movements that we perform are often determined not only by the spatial properties of the object's visible surfaces – in particular, their orientations and distances in depth – but also by the spatial properties of the object's currently occluded surfaces. As the French phenomenologist Merleau-Ponty might say, when we reach for an object, our grasping hand has 'knowledge' of the object's hidden features that our eyes do not have. This knowledge is displayed in our practical awareness of how to engage in actions that are kinematically appropriate to the object as a 3-D whole, and not just to its visible surface geometry.

Our ability to plan and engage in actions that successfully take account of an object's unseen features, I would now like to suggest, plays a major role in sustaining our sense that those features are really *present* though absent from sight. To borrow a nice example from Sean Kelly (2004), although we may not see the handle on the hidden side of a coffee mug, our awareness of the handle's spatial properties is revealed in how we act (and are prepared to act) in respect of the mug. An object's currently occluded

features may sometimes be as present in our 'behavioral space' – the space of possible bodily actions that are afforded us by the object – as are the object's currently visible features.¹⁸

This personal-level story comports well with behavioral and neuropsychological evidence concerning the types of neurally encoded information used in object-directed, visuomotor control. In particular, there is evidence that visuomotor actions targeted on a familiar object utilize both bottom-up, perceptual information about the object's visible, egocentric spatial properties as well as stored, categorical information about the object's perspective-invariant, volumetric shape (Milner and Goodale, 1995/2006; Jeannerod, 1997, 2006; Jacob and Jeannerod, 2003). When we act in respect of an object, the spatial parameters of the movements that we perform are thus adapted not only to the spatial properties of the object that are visible from our current perspective, but also to the spatial properties of the object that *would* be visible were we to see the side of the object facing away from us. This explains why we are usually able, for instance, to grasp a coffee mug by its handle, even when its handle is hidden at the time of action. Further details need not concern us here. The main point is that, in addition to NA-completion and C-completion, amodal perception plausibly also comprises what might be called *motoric completion*. An object's currently occluded features are motorically completed when they figure in the contents of 'action representations' (Jeannerod, 2006) that underwrite our abilities to interact with the object in ways that are appropriate to its 3-D, volumetric shape.¹⁹ In addition to NA-completion and C-completion, motoric completion plays an important role in conferring a sense of real, if unseen, presence on the occluded features of the objects that we perceive.²⁰

Department of Philosophy
Ohio University

NOTES

¹ For a review of findings and theoretical applications, see Kosslyn 1994 and Kosslyn *et al.* 2005.

² Dummett calls the exercise of this ability 'proto-thought' (1994, pp. 122–123), while Gosselin and Schyns (2003) refer to 'superstitious perception.'

³ Make-perceive also enables human beings to solve problems in ways that would not be possible on the basis of bottom-up perceiving or top-down imagining alone. Leon Battista Alberti advised painters to 'present the forms of things seen on the [frontal] plane as if it were of transparent glass' (1966, p. 51). In order to represent an object's 2-D, perspective projection, however, an artist need not use any of the mechanical devices contrived for this purpose during the Renaissance. With training, an artist can learn to notice an object's 2-D, perspective projection by viewing the object and 'tracing' its contour in imagination (Briscoe, 2008). Make-perceive has also been used for purposes of long-distance, non-instrumental

navigation. The Caroline islanders of Micronesia imaginatively construct linear constellations with fixed azimuths on the eastern and western horizons that convert the visible field of stars into a sidereal compass (Hutchins, 1995, ch. 2). This enables them to maintain accurate bearings of distant islands that may be well out of sight. Hutchins notes that ‘This seeing is not a passive perceptual process. Rather, it is the projection of external structure (the arrangement of stars in the heavens) and internal structure (the ability to identify the linear constellations) onto a single spatial image’ (1995, p. 68). Both of these examples illustrate the cognitively sophisticated purposes to which make-perceive can be put. For related examples, see Hutchins, 2005.

⁴ Three remarks: First, Nanay characterizes amodal perception as the representation of the non-visible *parts* of perceived objects. It is clear from Nanay’s discussion, however, that amodal perception also encompasses the representation of non-visible *properties* such as shape, color, and texture. Accordingly, I shall here use the more general term ‘feature’ where Nanay uses ‘part.’ Second, although Nanay focuses in his paper on the problem of how we represent the *occluded* features of perceived objects, he offers a more general definition of amodal perception: ‘We perceive a part of a (perceived) object amodally if we receive no sensory stimulation from that part of the object’ (p. 241). Like Nanay, I shall be focusing here on the problem of amodal perception as it arises in the case of occlusion. Last, projected mental imagery is supposed to account for the representation of occluded features both in cases of *superposition*, where visible surfaces hide detached surfaces in the more distant background, and *self-occlusion*, where an object’s near side occludes its far side. In the former case, both visible surfaces and hidden background surfaces are presumably represented from a single visual perspective or point of view. In the latter case, however, the perspective of visual perception and the perspective of visual imagination seem to come apart. The spatial point of view from which I see the visible surfaces of a house (and egocentrically locate them relative to myself) and the point of view from which I imagine its self-occluded surfaces, i.e. the surfaces I would see were I to view the house from a position facing its far side, are *different* points of view. Hence, it is not clear that the account Nanay provides of amodal perception in cases of superposition can be straightforwardly extended to amodal perception in cases of self-occlusion. Thanks to a referee for emphasizing the need to make the first two of these remarks.

⁵ The claim that ‘imagination is a necessary ingredient of perception itself’ originates with the theory of productive imagination developed in the *Critique of Pure Reason* (Kant, 1787/1999, A120). Nanay’s proposal that seeing an object involves projecting mental images of its currently occluded features has a clear, if unacknowledged, affinity with the reconstruction of the theory put forward by Wilfrid Sellars (1978/2007). According to Sellars, productive imagination is a capacity that enables us to form hybrid, ‘sense-image’ models of external objects in perceptual experience. The productive imagination performs this function in part by supplementing awareness of an object’s ‘occurrent sensible features’ with mental images of its hidden features. Sellars writes: ‘We do not see *of* the apple its opposite side, or its inside, or its internal whiteness . . . But while these features are not *seen*, they are not *merely* believed in. These features are present in the object of perception as actualities. They are present [in our visual experience] by virtue of being imagined’ (1978/2007, p. 458).

⁶ Noë has an even broader conception of amodal perception. According to Noë, amodal perception also includes shape- and size-constancy (2004, pp. 61–63; 2005, pp. 241–243). On this fairly radical view, we do not strictly speaking see that, e.g., a coin is objectively round or that a distant house is objectively bigger than a nearby mailbox.

⁷ Background knowledge, however, can *indirectly* affect perceptual grouping and segmentation operations in NA- and NM-completion by influencing which regions of the visual scene are attentionally selected (Pylyshyn, 1999; Prinz, 2000; Raftopoulos, 2009, ch. 7).

⁸ I should emphasize that this characterization is non-committal with respect to the question of whether other visual processes, e.g. those subserving color constancy, are cognitively penetrable.

⁹ Nanay emphasizes this point in the context of criticizing Noë's (2005) account of the 'perceptual presence' of the cat's hidden tail: 'According to [Noë's] account, what matters is my immediate perceptual access to the occluded object, and this access does not depend on whether I *know* how cat tails look. Thus, it follows from this view that I would have the same perceptual access to the cat's tail whether or not I know *how* cat tails look' (p. 247, my emphasis). Noë, however, is quite clear that only a perceiver with knowledge of the cat's sensorimotor profile, i.e. knowledge about the way the cat would visually appear were it viewed from different perspectives, could experience the cat's hidden tail as perceptually present.

¹⁰ Gibson, of course, would not talk of *representations* in the visual system, but this point can be put aside for present purposes.

¹¹ According to what vision scientists sometimes call 'law of generic views,' the visual system when confronted with more than one possible surface interpretation of a scene 'assumes' that the viewer is seeing the scene from a generic, i.e. a non-accidental, vantage point. The result is that the 'cap' on a T-junction is typically perceived as belonging to a surface in front and the 'stem' to a partially occluded surface that continues behind. For discussion, see Hoffman, 1998.

¹² Dretske (1981, 1988) suggests that a signal *r* carries the information that *s* is *F* if and only if the conditional probability that *s* is *F*, given *r* and background knowledge *k*, is 1 (but given *k* alone, less than 1). It is clear that no visual signal could carry information for occlusion given this strict requirement. The signal produced by Figure 4b above, e.g., is fully consistent with a visual scene containing eleven spatially separate surfaces (indeed with indefinitely many different visual scenes). It is also clear, however, that the visual system *does* form perceptual representations of occluded object features on the basis of information for occlusion. This, I take it, shows that the kind of 'natural information' actually exploited by the visual system depends on statistical frequencies in the environment as opposed to natural law or logical necessity. For related discussion, see Millikan, 2004, chs. 3–4.

¹³ This can be experienced more vividly with the stereoscopic variants presented in Anderson *et al.*, 2002, p. 154.

¹⁴ Nanay says that this is a 'new objection' to the belief-based account (p. 243), but it is actually long familiar from anti-empiricist/inferentialist approaches to perceptual organization in the Gestalt tradition. See, e.g., Koffka, 1935; Kanizsa, 1979; Kanizsa and Gerbino, 1982; and Pylyshyn, 2003, ch. 2.

¹⁵ Contrary to Nanay, there is substantial debate among early vision researchers about whether NM-completion and NA-completion are explained by identical neural mechanisms. One reason to think that they are not, as Anderson *et al.* (2002) and Singh (2004) point out, is that NM-completion and NA-completion occur under different stimulus conditions and generate contours with different shapes. For other critical assessments of the 'identity hypothesis,' see Fleming and Anderson, 2004 and Anderson, 2007a, b.

¹⁶ As mentioned above, there is evidence that NA-completed contours are represented in V1 (Sugita, 1999; Bakin *et al.*, 2000; von der Heydt, 2004). This by itself is sufficient to show that the belief-based account of NA-completion contradicts relevant empirical evidence, since the belief-based account would presumably predict that NA-completed contours are not represented in early visual processing areas.

¹⁷ There is no reason, aside from attentional constraints, to think that varieties of amodal perception distinguished here are mutually exclusive. For example, some of an object's

occluded features may be NA-completed, while others may be imagined or inferred to exist (Kanizsa and Gerbino, 1982).

¹⁸ This line of argument, richly articulated by Merleau-Ponty (1945/1962), originates within the Gestalt psychological tradition. See especially Koffka, 1935.

¹⁹ Noë (2004, 2005) has proposed that an object's occluded features are 'perceptually present' because we have practical knowledge of how proximal sensory stimulations caused by the object would vary as a consequence of movement. (For critical assessment, see Briscoe, 2008 and Leddington, 2009.) What I am here calling motoric completion does not depend on any such ability to anticipate the effects of movement on sensory stimulation. Rather, it depends on our ability to *interact* with the object in ways that take account of its currently unseen, spatial properties. The role that I am attributing to bodily action in amodal perception is thus quite different than that attributed to it by Noë and criticized by Nanay.

²⁰ I am grateful to Ruth Millikan, Lisa Mosier, and Jon Winawer for helpful discussion of this paper. I would also like to thank an audience at Auburn University, especially Keren Gorodeisky, Arata Hamawaki, and Kelly Jolley, as well as two anonymous referees for comments that resulted in significant improvements.

REFERENCES

- Alberti, L. (1966). *On Painting*. New Haven, CT: Yale University Press.
- Anderson, B. (2007a). 'The Demise of the Identity Hypothesis and the Insufficiency and Nonnecessity of Contour Relatability in Predicting Object Interpolation,' *Psychological Review* 114, pp. 470–487.
- Anderson, B. (2007b). 'Filling-in Models of Completion,' *Psychological Review* 114, pp. 470–487.
- Anderson, B., Singh, M. and Fleming, R. (2002). 'The Interpolation of Object and Surface Structure,' *Cognitive Psychology* 44, pp. 148–190.
- Bakin, J., Nakayama, K. and Gilbert, C. (2000). 'Visual Responses in Monkey Areas V1 and V2 to Three-Dimensional Surface Configurations,' *Journal of Neuroscience* 20, pp. 8188–8198.
- Briscoe, R. (2008). 'Vision, action and make-perceive,' *Mind and Language*, 23, 457–497.
- Dretske, F. (1981). *Knowledge and the Flow of Information*. Cambridge, MA: MIT Press.
- Dretske, F. (1988). *Explaining Behavior*. Cambridge, MA: MIT Press.
- Dummett, M. (1994). *Origins of Analytical Philosophy*. Cambridge, MA: Harvard University Press.
- Farah, M., Soso, M. and Dasheiff, R. (1992). 'Visual Angle of the Mind's Eye Before and After Unilateral Occipital Lobectomy,' *Journal of Experimental Psychology: Human Perception and Performance* 18, pp. 241–246.
- Fleming, R. and Anderson, B. (2004). 'The Perceptual Organization of Depth,' in L. Chalupa and J. Werner (eds) *The Visual Neurosciences*. Cambridge, MA: MIT Press, pp. 1284–1299.
- Gibson, J. J. (1966). *The Senses Considered as Perceptual Systems*. Boston, MA: Houghton Mifflin.
- Gibson, J. J. (1972). 'A Theory of Direct Visual Perception,' in J. Royce and W. Rozeboom (eds) *The Psychology of Knowing*. New York: Gordon and Breach, pp. 215–240.
- Gibson, J.J. (1979). *The Ecological Approach to Visual Perception*. Boston: Houghton Mifflin.
- Gosselin, F. and Schyns, P. (2003). 'Superstitious Perceptions Reveal Properties of Internal Representations,' *Psychological Science* 14, pp. 505–509.

- Grush, R. (2004). 'The Emulation Theory of Representation: Motor Control, Imagery and Perception,' *Behavioral and Brain Sciences* 27, pp. 377–442.
- Hoffman, D. (1998). *Visual Intelligence*. New York: W.W. Norton and Co.
- Hume, D. (1739/2000). *A Treatise of Human Nature*, D. Norton and M. Norton, eds. Oxford: Oxford University Press.
- Hutchins, E. (1995). *Cognition in the Wild*. Cambridge, MA: MIT Press.
- Hutchins, E. (2005). 'Material Anchors for Conceptual Blends,' *Journal of Pragmatics* 37, pp. 1555–1577.
- Jacob, P. and Jeannerod, M. (2003). *Ways of Seeing*. Oxford: Oxford University Press.
- Jeannerod, M. (1997). *The Cognitive Neuroscience of Action*. Oxford: Blackwell.
- Jeannerod, M. (2006). *Motor Cognition*. Oxford: Blackwell.
- Kamitani, Y. and Shimojo, S. (2004). 'Global Yet Early Processing of Visual Surfaces,' in L. Chalupa and J. Werner (eds) *The Visual Neurosciences*. Cambridge, MA: MIT Press, pp. 1129–1138.
- Kanizsa, G. (1979). *Organization in Vision*. New York: Praeger.
- Kanizsa, G. (1985). 'Seeing and Thinking,' *Acta Psychologica* 59, pp. 23–33.
- Kanizsa, G. and Gerbino, W. (1982). 'Amodal Completion: Seeing or Thinking?' in B. Beck (ed.) *Organization and Representation in Perception*. Hillsdale, NJ: Erlbaum.
- Kant, I. (1787/1999). *Critique of Pure Reason*. Cambridge: Cambridge University Press.
- Kellman, P. and Shipley, T. (1991). 'A Theory of Visual Interpolation in Object Perception,' *Cognitive Psychology* 23, pp. 141–221.
- Kelly, S. (2004). 'On Seeing Things in Merleau-Ponty,' in T. Carmon (ed.) *The Cambridge Companion to Merleau-Ponty*. Cambridge: Cambridge University Press.
- Koffka, K. (1935). *Principles of Gestalt Psychology*. London: Kegan Paul.
- Kosslyn, S. (1994). *Image and Brain: The Resolution of the Imagery Debate*. Cambridge, MA: MIT Press.
- Kosslyn, S. and Sussman, A. (1995). 'Roles of Imagery in Perception: Or, There is No Such Thing as Immaculate Perception,' in M. Gazzaniga (ed.) *The Cognitive Neurosciences*. Cambridge, MA: MIT Press, pp. 1032–1045.
- Kosslyn, S., Pascual-Leone, A., Felician, O., Camposano, S., Keenan, J., Thompson, W., Ganis, G., Sukel, K. and Alpert, N. (1999). 'The Role of Area 17 in Visual Imagery: Convergent Evidence from PET and rTMS,' *Science* 284, pp. 167–170.
- Kosslyn, S., Thompson, W. and Ganis, G. (2005). *The Case for Mental Imagery*. Oxford: Oxford University Press.
- Leddington, J. (2009). 'Perceptual Presence,' *Pacific Philosophical Quarterly* 90, pp. 482–502.
- Lee, T. and Nguyen, M. (2001). 'Dynamics of Subjective Contour Formation in the Early Visual Cortex,' *PNAS* 98, pp. 1907–1911.
- Lehar, S. (2003). *The World in Your Head*. Mahwah, NJ: Lawrence Erlbaum.
- Machery, E. (2009). *Doing Without Concepts*. Oxford: Oxford University Press.
- Merleau-Ponty, M. (1945/1962). *The Phenomenology of Perception*, C. Smith, trans. London: Routledge and Kegan Paul Ltd.
- Michotte, A., Thines, G. and Crabbe, G. (1964/1991). 'Amodal Completion of Perceptual Structures,' in G. Thines, A. Costall and G. Butterworth (eds) *Michotte's Experimental Phenomenology of Perception*. Hillsdale, NJ: Erlbaum, pp. 140–167.
- Millikan, R. (2004). *Varieties of Meaning*. Cambridge, MA: MIT Press.
- Milner, D. and Goodale, M. (1995/2006). *The Visual Brain in Action*, 2nd edn. Oxford: Oxford University Press.

- Nakayama, K., He, Z. and Shimojo, S. (1995). 'Visual Surface Representation: A Critical Link Between Lower-Level And Higher-Level Vision,' in S. Kosslyn and D. Osherson (eds) *Visual Cognition*. Cambridge, MA: MIT Press.
- Nanay, B. (2010). 'Perception and Imagination: Amodal Perception as Mental Imagery,' *Philosophical Studies* 150, pp. 239–254.
- Noë, A. (2004). *Action in Perception*. Cambridge, MA: MIT Press.
- Noë, A. (2005). 'Real Presence,' *Philosophical Topics* 33, pp. 235–264.
- Pessoa, L., Thompson, E. and Noë, A. (1998). 'Finding Out About Filling-In: A Guide to Perceptual Completion for Visual Science and the Philosophy of Perception,' *Behavioral and Brain Sciences* 21, pp. 723–802.
- Prinz, J. (2000). 'A Neurofunctional Theory of Visual Consciousness,' *Consciousness and Cognition* 9, pp. 243–259.
- Pylyshyn, Z. (1999). 'Is Vision Continuous with Cognition?: The Case for Cognitive Impenetrability of Visual Perception,' *The Behavioral and Brain Sciences* 22, pp. 341–365.
- Pylyshyn, Z. (2003). *Seeing and Visualizing: It's Not What You Think*. Cambridge, MA: MIT Press.
- Raftopoulos, A. (2009). *Cognition and Perception*. Cambridge, MA: MIT Press.
- Sellars, W. (1978/2007). 'The Role of the Imagination in Kant's Theory of Experience,' in K. Scharp and R. Brandom (eds.) *In the Space of Reasons: Selected Essays of Wilfrid Sellars*. Cambridge, MA: Harvard University Press, pp. 454–468.
- Singh, M. (2004). 'Modal and Amodal Completion Generate Different Shapes,' *Psychological Science* 15, pp. 454–459.
- Slotnick, S., Thompson, W. and Kosslyn, S. (2005). 'Visual Mental Imagery Induces Retinotopically Organized Activation of Early Visual Areas,' *Cerebral Cortex* 15, pp. 1570–1583.
- Sugita, Y. (1999). 'Grouping of Image Fragments in Primary Visual Cortex,' *Nature* 401, pp. 269–272.
- von der Heydt, R. (2004). 'Image Parsing Mechanisms of the Visual Cortex,' in L. Chalupa and J. Werner (eds) *The Visual Neurosciences*. Cambridge, MA: MIT Press, pp. 1139–1150.