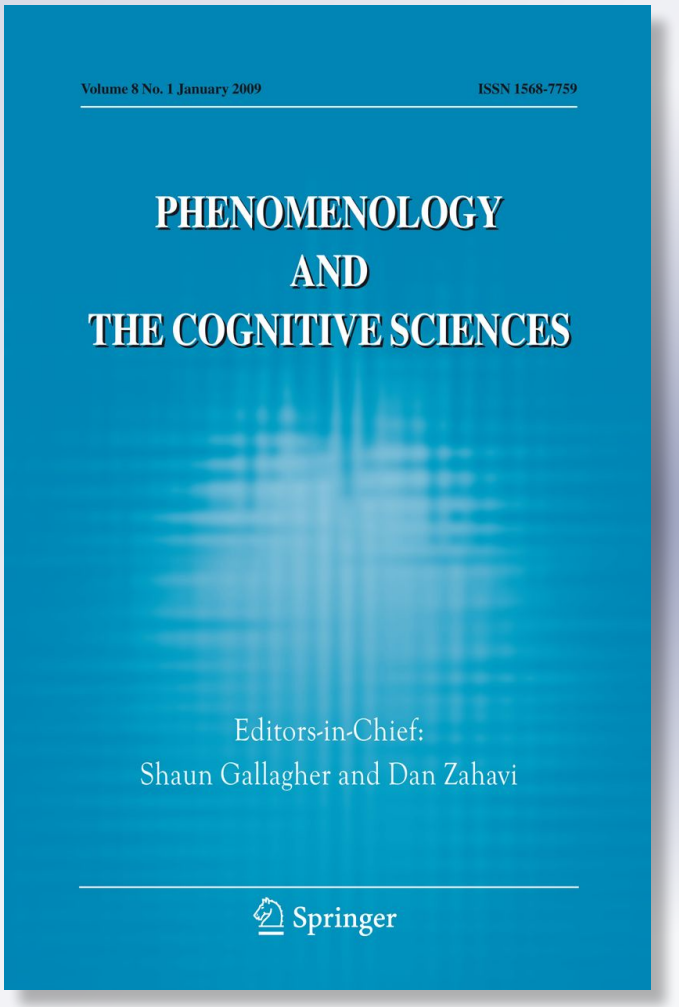
Perceiving pictures

Bence Nanay

Phenomenology and the Cognitive Sciences

ISSN 1568-7759
Volume 10
Number 4

Phenom Cogn Sci (2011) 10:461-480
DOI 10.1007/s11097-011-9219-x



Springer

Phenom Cogn Sci (2011) 10:461–480
DOI 10.1007/s11097-011-9219-x

Perceiving pictures

Bence Nanay

Published online: 26 August 2011

Abstract I aim to give a new account of picture perception: of the way our visual system functions when we see something in a picture. My argument relies on the functional distinction between the ventral and dorsal visual subsystems. I propose that it is constitutive of picture perception that our ventral subsystem attributes properties to the depicted scene, whereas our dorsal subsystem attributes properties to the picture surface. This duality elucidates Richard Wollheim's concept of the "twofoldness" of our experience of pictures: the "visual awareness not only of what is represented but also of the surface qualities of the representation." I argue for the following four claims: (a) the depicted scene is represented by ventral perception, (b) the depicted scene is not represented by dorsal perception, (c) the picture surface is represented by dorsal perception, and (d) the picture surface is not necessarily represented by ventral perception.

Keywords Picture perception · Dorsal stream · Ventral stream · Twofoldness

Picture perception and "twofoldness"

What perceptual state are we in when we see an object in a picture? This question is a crucial one both in the philosophy of perception and in aesthetics, and the aim of this paper is to outline a possible way of answering it.[1]

[1]There are also many ways of raising this question. Some talk about the *experience* of seeing something in the picture, some talk about the (not necessarily conscious) *perceptual state* of seeing something in a picture. It seems that we are capable of perceiving things in pictures unconsciously. We can perceive objects in pictures even if we are not conscious of either the surface or the depicted object, as the widely discussed phenomenon of subliminal priming shows (Strahan et al. 2002; Eimer and Schlaghecken 2003; Greenwald et al. 1996). Perception can be conscious and unconscious and so can perceiving something in a picture. Thus, it seems that perceiving something in a picture does not need to be a conscious experience. I will talk about cases where we consciously see an object in the picture. But as the account of picture perception I give in this paper does not make any reference to consciousness, it can also be applied in the case of unconscious picture perception.

B. Nanay (✉)
Department of Philosophy, University of Antwerp, D 413, Grote Kauwenberg 18, 2000 Antwerp, Belgium
e-mail: bn206@cam.ac.uk
URL: http://webh01.ua.ac.be/bence.nanay

B. Nanay
Peterhouse, University of Cambridge, Cambridge CB2 1RD, UK

Following Richard Wollheim, I will call the experience of seeing something in a picture seeing-in (Wollheim 1980, 1987, 1998). I will assume throughout the paper that our experience of seeing something in a picture is a *perceptual* experience. It has been suggested that it is, at least partially, constituted by an imaginative episode (Walton 1990, 2002; Maynard forthcoming). I will not evaluate, or argue against, this suggestion here, but simply assume that seeing-in *really is seeing* the depicted object in a certain way, namely, in the picture, rather than imagining it a certain way. In other words, I'm assuming that picture perception is indeed a perceptual process (I'll say a bit more about this assumption below).

The question about the nature of seeing-in is crucial enough, but it has also been suggested that seeing-in may play an important role in understanding what pictures are: pictures are those objects in the face of which "suitably informed" perceivers are supposed to have a seeing-in experience (Wollheim 1987, 1998). I will remain neutral about whether we can use seeing-in to define what pictures are. Even if one is persuaded by an account of depiction that makes no reference to seeing-in (as many in fact are, see Goodman 1968; Kulvicki 2006; Schier 1986, 1993; Peacocke 1987), it is still a crucial question what makes picture perception different from face to face perception.

So the question is what experience we are supposed to go through when we see a picture of an object. What happens in our mind when we see a depicted object in a picture? Ernst Gombrich claims that what is constitutive of this experience is that our attention alternates between the two-dimensional surface and the three-dimensional represented object (Gombrich 1960). Richard Wollheim, in contrast, argues that the experience we are supposed to go through when seeing pictures is a twofold one: we are simultaneously aware of the picture surface and the represented object (Wollheim 1980, 1987, 1998; Nanay 2004, 2005; see also Lopes 1996, 2005). As Wollheim puts it:

> The spectator is, and remains, visually aware not only of what is represented but also of the surface qualities of the representation. (Wollheim 1980, p. 214–215).

This feature of our experience of pictures is called "twofoldness" and in some form or other, many philosophical accounts of seeing-in endorsed it as a necessary feature of our experience of pictures (Walton 1990, pp. 300–301, Walton 2002, p. 33, and Walton 1991, p. 423; see Nanay 2004; Maynard 1994, esp. pp. 158–159 on the differences between Walton's and Wollheim's concept of twofoldness; see also Lopes 2005, chapter 1; Kulvicki 2006, pp. 172–173; Hopkins 1998, esp. pp. 15–17 for moderately critical overviews).[2]

In spite of the widespread use of this notion, neither Wollheim nor other philosophical accounts of pictorial perception say much about what is supposed to be meant by the twofoldness of experience. The account I argue for in this paper accepts Wollheim's original insights about the importance of twofoldness in our

[2] Gombrich's account of our experience of pictures is inconsistent with the idea of twofoldness. As he said: "is it possible to 'see' both the plane surface and the battle horse at the same time? If we have been right so far, the demand is for the impossible. To understand the battle horse is for a moment to disregard the plane surface. We cannot have it both ways…" (Gombrich 1960, p. 279.)

experience of perceiving pictures. If we want to understand seeing-in, we need to understand twofoldness. But there is some terminological confusion around the concept of twofoldness that needs to be cleared up.

Wollheim talks about simultaneous *awareness* of surface and scene (not just in the definition quoted above but also in Wollheim 1998, p. 221; Wollheim 1987, p. 46). But the notion of awareness he uses is ambiguous and as a result Wollheim's notion of twofoldness itself is also ambiguous (see Nanay 2005). And those who used, or criticized, the notion of twofoldness did not get rid of this ambiguity. Here are two possible interpretations of twofoldness (both of which we have good reasons to attribute to Wollheim):

1. We consciously attend both to the depicted object and to some properties of the surface.[3]
2. We represent both the depicted object and some of the properties of the picture surface (while we may or may not attend to the surface).

We have many reasons to reject the view that (1) is a necessary feature of seeing-in, but none of these reasons should persuade us to reject (2) as a necessary feature of seeing-in. I will go through some of the influential criticisms of the concept of twofoldness and of the claim that twofoldness is necessary for seeing-in and point out that although they are all valid arguments if we take twofoldness to mean (1), they lose their appeal if we interpret twofoldness as (2).

First, if twofoldness implies conscious attention[4] to surface and scene, then this rules out the possibility of seeing things in pictures unconsciously. But, as we have seen in footnote 1, it seems that we are capable of seeing things in pictures unconsciously: we can see objects in pictures even if we are not conscious of either the surface or the depicted object, for example, in the case of subliminal priming. Seeing can be conscious and unconscious and so can seeing-in. So, while we have good reason to reject (1) as a necessary condition for seeing-in, these considerations do not count against the claim that seeing-in entails (2), the simultaneous visual *representation* of the picture surface as well as the depicted object.

Second, Jerrold Levinson rightly pointed out that most of the time when we look at pictures (say, when we are looking at commercials or magazines) we *ignore* the surface: "it is far from clear that when you see the woman in the picture you must in some measure be attending to, taking notice of, or consciously focusing on the picture's surface or patterning as such" (Levinson 1998, p. 229; Lopes 1996, pp. 37–51 makes a similar point). Thus, if Wollheim held that (1) is a necessary feature of seeing-in, he was wrong. But Levinson's considerations are silent about whether (2) the other interpretation of twofoldness could be considered to be a necessary feature of seeing-in.

[3] One important consideration in favor of (1) is the following quote: "The seeing appropriate to representations permits simultaneous attention to what is represented and to the representation" (Wollheim 1980, p. 213). But it is not clear whether seeing-in only needs to "permit" simultaneous attention or it is constituted by it.

[4] I have been and I will be using the concepts of "attention" and "conscious attention" interchangeably. This does not mean that I take sides in the grand debate about whether attention is necessary and/or sufficient for consciousness. On this thorny issue, see Campbell (2002) and Prinz (2010).

Dominic Lopes gives a thorough analysis of the notion of twofoldness (Lopes 2005, chapter 1). He concludes that although twofoldness is an important feature of some of our experiences of some pictures, it is not a necessary feature of seeing-in in general. As he says, "It is only in virtue of seeing the configuration of marks on its surface […] that we see anything at all in the picture. However, seeing a pictorial design face to face does not entail seeing the design as a design" (Lopes 2005, p. 28, see also p. 35). Thus, we always see the surface of the picture when we see something in it, but we may or may not see its design as design. This sounds like a rejection of the claim that (1) is a necessary feature of seeing-in and an endorsement of the claim that (2) is. But as Lopes restricts the notion of twofoldness to (1) (see Lopes 2005, p. 33), he denies that twofoldness (that is (1)) is a necessary feature of seeing-in.

We have found that there are a number of reasons to reject the idea that (1) is necessary for seeing-in. But none of the arguments that have been given against the importance of twofoldness show that (2) is not necessary for seeing-in. I will assume in what follows that twofoldness is to be interpreted as (2) and that it is (2) that is taken to be a necessary feature of seeing-in. This is not such a wild assumption: many of the critics of the notion of twofoldness and of Wollheim's views in general would also be happy with this claim (Lopes 2005, chapter 1; maybe even Hopkins 2010, §1). Nor is it a very informative assumption as depending on how exactly we think of twofoldness, our account of picture perception will be different.

My aim in this paper is to give a new account of picture perception, which could be thought of as an extension of Wollheim's account in as much as I also take twofoldness, that is, (2), to be necessary for the experience we are supposed to go through when we are looking at pictures. But in order to do so, I aim to reinterpret, and give substance to, the Wollheimian notion of twofoldness. My claim will be that the twofold experience of pictures Wollheim talks about corresponds to the dichotomy between our dorsal visual processing of the surface of the picture and our ventral visual processing of the depicted scene.

Picture perception and the two visual subsystems

Humans (and other mammals) have two visual subsystems that use different regions of our central nervous system, the ventral and dorsal streams. These two streams can be differentiated anatomically and functionally. I will stick to the latter way of drawing this distinction. To put it very simply, the ventral stream is responsible for identification and recognition, whereas the function of the dorsal stream is the visual control of our motor actions. In normal circumstances, these two systems function together, but if one of them is removed or malfunctioning, the other can still function relatively well (see Milner and Goodale 1995; Goodale and Milner 2004; Jeannerod 1997, for overview).

If the dorsal stream is malfunctioning, the agent can recognize the objects in front of her, but she is incapable of manipulating them or even localizing them in her egocentric space (especially if the perceived object is outside the agent's fovea). This happens if a patient is suffering *optic ataxia*. If the ventral stream is malfunctioning, the agent can perform actions with objects in front of her relatively well, but she is

incapable of even guessing what these objects are. This happens in the case of *visual agnosia*.

The philosophical implications of this physiological distinction are not at all clear. Some argued that ventral visual processing is conscious whereas dorsal is unconscious (see esp. Milner and Goodale 1995; Goodale and Milner 2004), but this view has been criticized both on empirical and on conceptual grounds (see for example, Dehaene et al. 1998; Jeannerod 1997; see also Jacob and Jeannerod 2003 for a summary). It has also been suggested that dorsal processing gives rise to nonconceptual content, whereas ventral processing gives rise to conceptual content (see Clark 2001 for a summary on the literature on this). I do not need to take sides in either of these questions.

All I assume in this paper is that these two streams of visual processing are *functionally* different—I want to remain neutral about whether underlying anatomical structures are separate (but see Jeannerod 1997 for a summary of the evidence of rich interactions between the ventral and the dorsal regions). The dorsal subsystem feeds into our perceptually guided actions and provides visual control for our motor actions, thus, it represents the perceived object egocentrically (more on what this may mean below). The ventral subsystem, however, feeds into our epistemic apparatus and helps us to recognize objects. This way of drawing the distinction between the two visual subsystems seems to be the common denominator between various interpretations of the anatomical findings: a similar distinction is made by Jeannerod (1997), Jacob and Jeannerod (2003), Goodale and Milner (2004), and Matthen (2005)—who labels the two as "motion-guiding vision" (dorsal) and "descriptive vision" (ventral). I will keep using the terms "ventral" and "dorsal" to refer to these two visual subsystems.

In healthy humans, the way the dorsal and the ventral stream works can come apart in some circumstances, as in the case of the three-dimensional Ebbinghaus illusion. The two-dimensional Ebbinghaus illusion is a simple optical illusion. A circle that is surrounded by smaller circles looks larger than a circle of the same size that is surrounded by larger circles. The three-dimensional Ebbinghaus illusion reproduces this illusion in space: a poker chip surrounded by smaller poker chips appears to be larger than a poker chip of the same diameter surrounded by larger ones. The surprising finding is that although our judgment of the comparative size of these two chips is wrong as we judge the first chip to be larger than the second one, if we are asked to pick up one of the chips, our grip-size is not influenced by the illusion (Aglioti et al. 1995, but cf. Gillam 1998 as well as Franz et al. 2003). The usual way of explaining this finding is that our dorsal stream is not fooled by the illusion but our ventral stream is.

The same results can be reproduced in the case of other optical illusions. In the Müller-Lyer illusion, while we (mistakenly) see the two lines as having different length, our eye and pointing movements represent them (correctly) as being the same (Goodale and Humphrey 1998; Gentilucci et al. 1996; Daprati and Gentilucci 1997; Bruno 2001).[5] Similarly, in the case of the "Kanizsa compression

[5] These results have been questioned as they are difficult to parse with the phenomenon that we also experience the Müller-Lyer illusion in the haptic sense modality (see Heller et al. 2002; Heller et al. 2005; Suzuki and Arashida 1992).

illusion" and the "hollow face illusion," our perception is deceived but our action is not (Bruno and Bernardis 2002; Króliczak et al. 2006, respectively). Thus, it does happen under exceptional circumstances that our ventral visual subsystem attributes a different property to an object from the one the dorsal subsystem does.

My claim is that our visual system functions in a somewhat similar manner when we are perceiving pictures. I will argue that the dorsal and the ventral visual subsystems attribute different properties to the perceived object whenever we see objects in pictures. The ventral subsystem attributes properties to the depicted scene whereas the dorsal subsystem attributes properties to the surface of the pictures.

Or, to put it very simply, it is constitutive of our experience of seeing things in pictures that the depicted scene is represented by our ventral vision, whereas the surface of the picture is represented by our dorsal vision. Although this latter formulation is simpler, it is also slightly ambiguous, but this ambiguity can be easily resolved.

Some depicted objects do not exist. As a result, according to most philosophical analyses of perception, we do not literally perceive these objects as we cannot perceive something that does not exist (Clarke 1965). What is important for our purposes is that even if we do not (literally) perceive depicted objects, we are in perceptual states that represent these objects (whether or not they exist) as having certain properties. To sum up, the ventral subsystem represents the depicted scene (whether or not it exists) as having certain properties, whereas the dorsal subsystem represents the picture surface as having certain properties.

Clarification: the main claims of this paper are formulated in terms of perceptual representation: when we see things in pictures, the ventral subsystem represents the depicted scene, whereas the dorsal one represents the picture surface. But it has been argued that "perception is not a process of constructing internal representations" (Noë 2004, p. 178; see also Campbell 2002; Travis 2004; Martin 2006; Brewer 2006; Ballard 1996; Brooks 1991; O'Regan 1992; Hutto and Myin 2011; Chemero 2009). There are different versions of this claim (enactivism, relationalism, etc.), but it is important to emphasize that everything I say in this paper is consistent with both a representationalist and an anti-representationalist view of perception. For the sake of simplicity, I will talk about perceptual states representing various properties, but the equivalent claims can be made in anti-representationalist terminology (perceptual states "presenting" or "being sensitive to" or "tracking" some properties).

It is important, however, that when I say that the ventral or the dorsal subsystem represents certain properties (or that it attributes some properties), this claim is *not* about perceptual consciousness. Whether perceptual processes in the dorsal and the ventral subsystems are (necessarily? or necessarily not?) conscious is, as we have seen above, not an easy question to answer and nothing I say in this paper commits me one way or another. Similarly, what I mean by "ventral perception" and "dorsal perception" is to be understood accordingly: ventral perception is defined without any reference to perceptual consciousness: it is the perceptual process that is implemented by the ventral subsystem. The same goes for dorsal perception.

The argument

I need to argue for four claims in order to show that the ventral subsystem represents properties of the scene, whereas the dorsal one represents properties of the surface:

(a) The depicted object is represented by ventral perception
(b) The depicted object is not represented by dorsal perception
(c) The picture surface is represented by dorsal perception
(d) The picture surface is not necessarily represented by ventral perception

I will address these claims in turn.

(a) The depicted object is represented by ventral perception

I have assumed from the beginning that the experience of seeing objects in pictures is a genuinely perceptual experience: we really do see the depicted object in the picture. So the depicted object is represented perceptually: the question is whether it is represented by the ventral or the dorsal subsystem.

Before turning to this question, it is important to clarify that the assumption I am starting out with, namely, the claim that the depicted object is represented perceptually is not a particularly strong one. There has been a lot of discussion about what properties are represented perceptually (see Siegel 2006; Nanay 2011a; 2011b, forthcoming a, forthcoming b, for example). We perceive objects as having various properties. We perceive them as having a certain color, a certain shape and a certain size, for example. The question is whether we perceive them as having other kinds of properties that are less obviously perceptual, for example, the property of being a table. When I am looking at the object in front of me do I perceive it as a table or do I perceive it as having a certain shape, size and color and I only *infer* that it is a table?

I do not need to take sides in this debate, but I do need to point out that the assumption I made about the depicted object being represented perceptually is orthogonal to this debate. Maybe it is true that sortal properties are not represented in perception. The depicted object may still be represented perceptually, just as any other objects we see. Some properties of this object (say, its shape, size, and color) we represent perceptually, some others we don't. It may be true that some properties of the depicted object are inferred post-perceptually, but it does not follow that the depicted object itself is not represented perceptually. So claiming that the depicted object is represented perceptually is not such a wild assumption. And it does justice to the intuition that when we say that we see an object in a picture, it is not just a manner of speaking: it is indeed seeing the object in a certain way (see also Wollheim 1980, 1987; Matthen 2005, p. 308—but see Goodman 1968 for a contrasting view).

What I need to show is that if the depicted object is represented perceptually, it is represented by the ventral visual subsystem. Here is why. First, when we look at pictures, we can recognize what they are of. Thus, the perceptual representation of the depicted object feeds into our recognitional apparatus—which is exactly what the ventral subsystem is for.

Second, an empirical argument for the claim that the depicted object is represented by the ventral subsystem is provided by the fact that patients suffering

visual agnosia are incapable of seeing objects in pictures (Turnbull et al. 2004; Westwood et al. 2002). These patients have a functioning dorsal stream but there are serious impairments in the ventral stream. D. M., one such patient, can copy the two-dimensional lines of a picture but she cannot tell whether these two-dimensional lines depict an impossible object. Further, she is not subject to optical illusions (Müller-Lyer, Ponzo) that are usually taken to presuppose our ability to see three-dimensional objects in two-dimensional figures (Turnbull et al. 2004).

(b) The depicted object is not represented by dorsal perception

The main claim of this subsection is that depicted objects are not represented by dorsal perception. Dorsal perception is what allows us to localize objects in our egocentric space and helps us perform actions with them. As the space represented by the picture is not our egocentric space, we cannot localize the depicted objects (that are in the space of the picture) in our egocentric space. As Matthen argues, "a picture gives you no information of the location relative to yourself. Suppose you are looking at a picture of two men shaking hands. *Where* are they? As far as what you can tell by seeing-in the picture, the question has no answer" (Matthen 2005, p. 315).

Some further care is needed to make this point.[6] How should we interpret what is meant by "the ability to localize an object in one's egocentric space"?[7]

One possibility would be to say that this ability is just the ability to interact with this object. This is, for example, the way Gareth Evans understood egocentric space, as he argued that egocentric space is action space (Evans 1982, 1985). This is also what Poincaré means when he writes that "[…] to localize an object simply means to represent to oneself the movements that would be necessary to reach it" (Poincaré 1905/1958, p. 47).

If we accept this interpretation of the ability to localize objects in one's egocentric space, then we have a neat argument for (b). We cannot touch, smell or grasp the depicted objects: we cannot perform actions with them. We can touch, smell or grasp the part of the surface that represents these objects, but not the depicted objects themselves. Thus, as dorsal vision is what allows us to perform perceptually guided actions, the depicted objects are not represented dorsally (Matthen 2005, p. 312).

However, it could be argued that the ability to manipulate an object physically is not necessary for having the ability of localizing it in one's egocentric space. If I see an apple in a thick plexiglass container in front of me, I cannot touch it or grasp it or

[6] It is important to note that the question I raise here is different from the question about whether egocentric localization is a necessary condition for seeing per se (Walton 1984, 1997; Currie 1995, chapter 2, esp. p. 70; Currie 1991; Carroll 1995, p. 71; see also Warburton 1988; Cohen and Meskin 2008; Nanay 2010b). Even if egocentric localization is not necessary for seeing per se, it is still necessary for dorsal perception.

[7] It is important to distinguish this question from one about how some pictures locate the spectator at a very specific position in relation to the depicted scene. Although this line of thought has often been made in the psychoanalytic tradition of art and film theory (especially in connection with the concept of suture), it has also been discussed in the analytic tradition (for example, in chapter 3 of Wollheim 1987; see also Hopkins 2002). The question I am addressing here is not about the localization of the spectator as dictated by the pictorial space but about the localization of the depicted objects by the spectator.

manipulate it in any other way; still, I can localize it in my egocentric space. Further, we can point at depicted objects and pointing at an object is an action. Hence, there are at least some ways we can perform actions on the depicted object – but then we cannot use this way of conceiving of egocentric localization to argue for (b). Because of these considerations, I will not defend this rather demanding interpretation of localization in one's egocentric space, I would like to appeal to a less demanding notion of egocentric localization that is also less problematic.

The second, weaker, proposal for cashing out what localization in one's egocentric space means is the following. What is necessary for having the ability to localize an object in one's egocentric space, is that the agent represents the distance between the perceived object and herself in some (not necessarily conscious or explicit) way and she also has expectations about how this distance could be changed by moving closer to, or away from, the object. Thus, if I can localize an apple in my egocentric space, then if I move towards it (and the object itself does not move) than I should expect that the object will be closer to me than before (this ability is sometimes referred to as "sensorimotor skill," see Noë 2002, 2003, 2004; Campbell 2011; Nanay 2010b).[8]

It is important that representing the distance between the perceived object and oneself and having expectations about how this distance could be changed by moving closer to, or away from, it are supposed to provide only a necessary condition for localization in one's egocentric space. I do not need to give a full account of egocentricity here. All I need in order to argue for (b) is to show that a necessary condition for localization in one's egocentric space is not satisfied in the case of the perception of a depicted object.[9]

Even if we accept this weaker notion of what it means to be able to localize an object in one's egocentric space, we can still establish (b). If we move towards the direction of a depicted apple, we do not thereby move closer to the apple: we move closer to the picture surface, but not the apple.

An important consequence of both ways of understanding localization in one's egocentric space needs to be made explicit. Regardless of whether we accept the

[8] Noë emphasizes that we can localize even those objects in our egocentric space that we cannot manipulate or physically interact with, as long as we have expectations with regards to how they would change if we moved closer to it—as in the case of the object in the thick plexiglass container (see Noë 2004, chapter 3).

[9] A possible worry about using either of the above notions of localization in egocentric space in order to argue for (b) is the following. Suppose that I'm looking at myself on CCTV and there is an apple in front of me. I can see the apple on the screen and I can also touch it. Further, I do represent the distance between the apple and myself. Doesn't this mean that I can localize the apple in my egocentric space? If so, then it seems that the argument for (b) does not go through: we dorsally represent the depicted object. This is a complex and interesting case and I cannot address it in sufficient detail here. Note, however, that although I can touch the apple in front of me, I cannot touch the apple that is on the screen. More importantly, although I represent the distance between the apple in front of me and myself, I do not represent the distance between the apple I see on the screen and myself. If I see the apple on the screen, then the object of my perception is the apple on the screen and not the apple in front of me and what matters for egocentric localization is whether I represent the distance between myself and the object of my perception. And in this example, I do not represent the distance between the object of my perception and myself. I may represent the distance between the apple in front of me and myself (as we are in the same space) and I may also represent the distance between the depicted apple and the depicted myself (as they are in the same space), but I do not represent the distance between the depicted apple and myself. Hence, I do not localize it in my egocentric space.

weaker or the stronger notion of localization in one's egocentric space, it seems to follow that distant objects, like a mountain range in the distance are not dorsally perceived. We cannot touch, smell or grasp these distant objects: we cannot perform immediate physical actions with them. As dorsal vision is what allows us to perform perceptually guided actions, we have no reason to suppose that these objects are not represented dorsally. But the perception of distant objects will not count as dorsal perception according to the second, weaker, account of egocentricity either: we have no, or only very vague, (egocentric) representation of the distance between these distant objects and us and we have no (or only very vague) expectations about how this distance could be changed by moving closer to, or away from, them. Is this similarity between the perception of depicted objects and the perception of distant objects particularly surprising or counterintuitive? I don't think so: Matthen (2005, pp. 322–323), for example, explicitly and happily endorses it. I will return to this consequence of explaining egocentricity in the next subsection.[10]

(c) The surface of the picture is represented by dorsal perception

Suppose I am looking at a postcard. If someone asks me to touch the surface of the picture, I would have no problem doing so. If someone asked me to move in such a way that I see the postcard head on, I can do that easily. As dorsal perception is supposed to guide these actions, it seems that we must represent the surface dorsally. But let us proceed more slowly.

It is not a particularly controversial claim that we visually represent the surface when we recognize the depicted object in a picture.[11] It has been shown that there is a significant difference between our judgment of the size of a perceived object if this object is depicted (even in a hyper-realistic way) and our judgment when we see the same object through a glass, screen or colored glass. This phenomenon is taken to demonstrate that we represent the surface of the picture perceptually (see Hagen et al. 1978).

The controversial question is, of course, *in what way* we represent the surface: ventrally, dorsally or maybe both ventrally and dorsally? In order to settle this

[10] It has been argued that instead of two visual subsystems, we need to talk about three: the ventral, the ventrodorsal and the dorsodorsal. Thus, what has been taken to be one single dorsal subsystem should be divided into two: one responsible for manipulating objects (dorsodorsal) and one responsible for localizing in egocentric space (ventrodorsal) (Rizzolatti and Matelli 2003). My strategy was to show that we do not localize depicted objects in our egocentric space. Thus, the argument I presented in this section, rephrased using the terminology of the three visual subsystems framework, aimed to show that our *ventrodorsal visual subsystem* does not represent the depicted object.

[11] It is interesting to note that there are occasions where we do not represent the picture surface. Taking aside various anecdotes of insects that tried to fly through the canvas of a still-life, the obvious examples are *trompe l'oeil* paintings. If we are genuinely fooled by the *trompe l'oeil*, then we would readily try to reach through the canvas—we would not be aware of the surface. It is a controversial question whether *trompe l'oeil* paintings count as genuine instances of depiction (Wollheim 1987; Wollheim 1998; Levinson 1998; Feagin 1998; Lopes 2005). Whether or not they do, seeing *trompe l'oeil* paintings (that is, the experience of being fooled by them, see Lopes 2005 for a detailed analysis of our experience of *trompe l'oeil* paintings) is not a twofold experience, as we are supposed to be unaware of the fact that it is a picture at all that we are looking at. In other words, being fooled by *trompe l'oeil* paintings is not seeing-in (but this does not imply that *trompe l'oeil* paintings are not pictures).

question, I would like to appeal to a very widely discussed phenomenon in connection with picture perception.

The phenomenon is that if our position changes in front of the picture, our view of the depicted object does not change (Vishwanath et al. 2005; Pirenne 1970; Polanyi 1970; Wollheim 1980, pp. 215–216, Matthen 2005, pp. 315–317). Even if we look at a picture from an oblique angle, we don't see it as distorted. This is surprising and needs to be explained, as the projection of the depicted object on our retina is very different from the way it is when we look at the picture head on. The standard way of explaining this phenomenon is to say that we are perceptually aware of the orientation of the picture surface and this awareness compensates for the oblique view: that is why we do not see the depicted object as distorted. This proposal goes back at least as far as Pirenne's analysis allegedly inspired by a letter by Albert Einstein (see Pirenne 1970, pp. 99f).

What is interesting from our point of view is that there are cases where there is no such compensation. When we are looking at ceiling frescos from an oblique angle, for example, we do see the depicted scene as distorted. So what is the difference? Pirenne's original suggestion is that we do not have perceptual access to the orientation of the surface of the fresco, because it is too far away. When (because of the crowd) we are looking at the Mona Lisa from an oblique angle, however, we do have perceptual access to the orientation of the picture surface, which allows our perceptual system to compensate for the oblique view: our experience of the depicted scene is not distorted. Any explanation for the oblique perception of pictures needs to be able to tell not only why our perception of the Mona Lisa is not distorted but also why our perception of the distant fresco is distorted.[12]

So far, I have pretended that Pirenne's analysis of the "compensation" for the oblique point of view is uncontroversial. It is not. If Pirenne were right, then perceptual access to the orientation of the surface would be necessary and sufficient for compensating for the oblique angle and thus, for not experiencing the depicted scene in a distorted manner. But both the necessity and the sufficiency claims have been questioned. It seems that even if all the cues that indicate the orientation of the picture surface are artificially removed, we still experience the depicted scene without any distortions (Busey et al. 1990). Furthermore, it has also been argued that even if we do have cues that indicate the orientation of the picture surface, we sometimes do experience a distorted depicted scene (Halloran 1989). It needs to be noted that these two experiments are not considered conclusive (see Busey et al. 1990; Halloran 1989; Rogers 1995; Topper 2000; Kulvicki 2006; Koenderink et al. 2004). As Koenderink et al. 2004, p. 526, say "there appears to be some (weak) consensus that no 'correction' is applied to pictorial space due to obliquely viewed pictures" (see also Maynard 1996, p. 33).

But the "compensation" view, and the possibility of explaining what we experience when we look at pictures from an oblique angle, could be salvaged if

[12] An interesting related phenomenon is anamorphosis: when a surface, or part of a surface, is designed in such a way that one can see something in it only from a specific angle—as in the case of Holbein's *Ambassadors*, where the skull can only be seen from a very oblique angle. Note that the problem I am discussing here is importantly different from the problem of anamorphosis. In the case of anamorphosis, the depicted object can *only* be seen from an oblique angle, whereas the Pirenne problem is about how it is possible that the depicted object that can be seen head on can also be seen from an oblique angle.

we introduce a distinction between having ventral perceptual access to the orientation of the picture surface and having dorsal perceptual access to it. We need to differentiate between two versions of the "compensation" view: (i) our ventral representation of the orientation of the picture surface compensates for the oblique view and (ii) our dorsal representation of the orientation of the picture surface compensates for the oblique view.

Dorsal and ventral representation of the orientation of the picture surface comes apart in the usual ways: the latter feeds into our recognitional abilities whereas the former helps us to localize objects in our egocentric space and to interact with them. The arguments against the "compensation" view I quoted above are arguments against (i): against the claim that "compensation" entails *ventral* access to the orientation of the picture surface. To use just one example, in the experiment that is supposed to show that we do "compensate" even without perceptual cues about the orientation of the picture surface, these cues are cues that are ventrally represented (the "double projection technique" that Busey et al. (1990) use for removing these cues would remove ventral cues only: cues the perception of which feeds into our recognitional apparatus, see the discussion in Cutting 1987; Busey et al. 1990, p. 2). Thus, what this experiment really shows is that our ventral perceptual access to the orientation of the picture surface is not necessary for experiencing the depicted scene without any distortion.

My proposal is that we should reject (i) and accept (ii): we should interpret our perceptual access to the orientation of the picture surface as a dorsal phenomenon: if we do so, we do not face any of the objections outlined above and we can indeed use Pirenne's original observations to explain what we experience when we look at pictures from an oblique angle.

But remember that we are supposed to explain not only why our perception of pictures is not distorted when we are looking at pictures from an oblique angle. We are also supposed to explain why and when our perception of pictures *is* distorted. To go back to Pirenne's original fresco example, when we are looking at a ceiling fresco from an angle, we do experience the depicted scene as distorted. What is the difference? If we accept (ii), the hypothesis that our awareness of the orientation of picture surfaces is dorsal awareness, then we have a straightforward explanation. When we are looking at a ceiling fresco from an angle, we are lacking *dorsal* access to the orientation of the fresco – the fresco is too far away for our dorsal subsystem to allow localization in our (egocentric) space: we have no (egocentric) representation of the distance between the picture surface and us and we have no expectation about how this distance could be changed by moving closer to, or away from, it.

Some empirical studies also seem to support this hypothesis. Vishwanath et al. (2005) argued for a version of the "compensation" view, where they describe our perceptual access to the orientation of the picture surface as access to the "local slant" of various points of interest on the picture surface. Although the authors do not raise the question whether "local slant" is dorsally or ventrally represented, the fact that the representation of the "local slant" of one point of the surface is insensitive to (or, as they put it, "not contaminated by") both picture content and the "local slant" of other points of the surface suggests that it is dorsally represented. Why? Because dorsal (but not ventral) vision is taken to attribute properties

"locally," that is, in a way that is insensitive to the properties attributed to other parts of the perceived scene, as the experiments on optical illusions I mentioned below show (see also Jeannerod 1997; Goodale and Milner 2004).

Thus, we have good reason to suppose that our perceptual access to the orientation of the picture surface is dorsal. If we take our perceptual access to the picture surface to be dorsal, we can explain both why oblique view of pictures is not distorted and why under some circumstances (when the surface is too far away) it is distorted. If we take it to be ventral access, both of these explanations become problematic. Thus, we have good reason to accept (ii) over (i). And this is exactly what we need in order to support claim (c): the picture surface is represented not by the ventral but by the dorsal subsystem.

(d) The surface of the picture is not necessarily represented by ventral perception

It is crucial to note that the scope of this claim is different from that of the previous three. I will not argue that the surface is never represented by ventral perception, but only that it is not necessary for perceiving pictures that we represent the surface ventrally.

It would be easy to argue for (d) if we assumed that ventral perception is always conscious or that it always involves conceptualization. We can see an object in a picture even if we are not conscious of any of the surface properties and even if we are not conceptualizing any of the surface properties. But, as we have seen below both of these assumptions about ventral perception are problematic. It seems that the only uncontroversial claim about ventral perception is that it feeds into our recognitional abilities.

Take the following example. I am watching a soccer game, and see the penalty area, a simple rectangle, on the field. Because of the position of the camera, this rectangle projects as a trapezoid shape on the screen. Now, I do not need to, and I usually do not, recognize or identify a trapezoid shape on the screen when I see the penalty area on the screen.[13] But if someone asks me about the shape on the picture surface that depicts the penalty area, I would need to *actively direct my attention* to the surface properties in order to respond (see Clark 1960, pp. 17 and 36–37). Now, if the surface properties *were* represented ventrally, that is, in such a way that their representation would feed into our ability to identify and recognize, there would be no need for such active and intentional shift of attention—the ventral representation our recognitional apparatus needs to use would be there all along. As there *is* a shift of attention in these cases, it seems that the surface properties of pictures are not represented ventrally. Thus, most of the time when we see things in pictures we can do so without representing the surface ventrally.

But there are cases where we might do just this: where we do represent the picture surface ventrally, especially if we are interested in the way the coin (or the penalty area) is depicted. In this case, our attention may be drawn to some of the features of the surface: brushstrokes, composition, and so on. My point is that this is not

[13] This, in itself, says nothing about (d), as ventral representation is not identical to recognition, but only feeds into the latter.

necessary for seeing things in pictures. We can see an apple in a picture even if we do not ventrally represent any features of the surface.

It is an extremely interesting question *when* we represent surface features ventrally and the answer to this question may be a step towards understanding not the perception, but the aesthetic appreciation of pictures (Podro 1991, p. 173, Podro 1998, p. 28; Lopes 2005 pp. 40 and 128–9; Hopkins 2010; Nanay 2010a). It has been suggested that conscious attention to, and recognition of, the picture surface/design is necessary for the aesthetic appreciation of pictures (Podro 1991; Lopes 2005; and arguably, see Nanay 2005, this is also Wollheim's view in Wollheim 1980, 1987). If we add the not very strong assumption (an assumption I explicitly resist) that conscious attention and recognition implies ventral processing, then it is a short and easy step to conclude that ventral perception of the picture surface/design is necessary for the aesthetic appreciation of pictures. Let us suppose that those who argued that attention to, and recognition of, the properties of the surface is necessary for the aesthetic appreciation of pictures are right. And let us also suppose that conscious attention and recognition implies ventral processing, an assumption I have resisted because it is controversial, but I have not argued against it either (again, because it is controversial). Is (d) jeopardized if we put these two claims together? No it isn't. If we put these two claims together, what we get is that ventral representation of the surface is necessary for the aesthetic appreciation of pictures. But it does not follow from this that ventral representation of the picture is necessary *for picture perception*. The aesthetic appreciation of pictures is not necessary for picture perception: as we have seen in the first section, we can (and very often do) perceive pictures without appreciating them aesthetically (Levinson 1998; Lopes 1996, pp. 37–51). Hence, even if it is a necessary condition for the aesthetic appreciation of pictures that the surface engages our ventral vision, this claim should not be extended to picture perception in general.

Some empirical support

I argued that both the ventral and the dorsal subsystems are needed for the perception of objects in pictures. If this claim is to be taken seriously as a philosophical claim, it needs to be consistent with the empirical findings about patients whose dorsal or ventral visual subsystem is malfunctioning. In fact, my proposal implies that breakdown of either the ventral or the dorsal stream would lead to a breakdown in picture perception. And this is in fact what the empirical findings suggest.

There seems to be plenty of evidence for the claim that the malfunctioning of the ventral stream leads to a breakdown in picture perception. Patients with visual agnosia, as we have seen in the case of D.M., are extremely bad at picture perception (see Turnbull et al. 2004; Westwood et al. 2002).

But what happens if the dorsal stream is malfunctioning? In my account both the ventral and the dorsal subsystems are needed for picture perception. Thus, one would expect that if either of these two subsystems breaks down, we can no longer see things in pictures. This is exactly what happens in the case of visual agnosia, that is, if the ventral system breaks down. But can we make the parallel claim about optic

ataxia, that is, when the dorsal system malfunctions? The picture perception of patients with optic ataxia has not been tested thoroughly (David Milner, personal communication and Christopher Striemer, personal communication).

Furthermore, patients with optic ataxia tend to cope well with their environment, and the lack of the dorsal stream is only manifest under some special circumstances. For example, they tend to have no problem reaching for and grasping objects in their fovea, difficulties with manipulation only occur if they perceive something outside of their fovea. Also, as these patients have been growing up in a world full of pictures, it seems unlikely that they wouldn't acquire a non-dorsal way of recognizing *that* they perceive a picture. The picture perception of optic ataxia patients may be more difficult to assess than that of visual agnosia patients.

But there is one empirical study I know of that that could be considered to be decisive from the point of view of my proposal. A patient presenting symptoms of optic ataxia, A.T., who sustained a bilateral parieto-occipital infarct during eclampsia did perceive pictures, but her "evaluation of line length and size of drawn figures was poor" (Jeannerod et al. 1994, p. 370; see also Jeannerod 1997, p. 62). What we have here is a malfunctioning of picture perception as a result of a malfunctioning of the dorsal stream. The malfunctioning of the dorsal stream does not result in the complete breakdown of picture perception (like the malfunctioning of the ventral stream does), but it does lead to misestimating the distances and size of the depicted scenes—a phenomenon I address in the next section in a different context. In short, my proposal is consistent with, and even predicts, the relevant empirical findings.

A final objection

Here is a possible objection that could jeopardize the account of picture perception I have put forward. How would my account explain the fact that we can perceive pictures of pictures (Lopes 1996; Kulvicki 2006; Newall 2003)? I have argued (see (b) above) that our dorsal vision does not attribute properties to depicted objects. But then our dorsal vision cannot attribute properties to the surface of the depicted picture either. But as (see (c)) it is necessary for seeing things in pictures that our dorsal vision attributes properties to the picture surface, we cannot see anything in the depicted picture. But this does not sound right.

To put it briefly, suppose that I am looking at a picture X that depicts another picture Y depicting object Z. Because of (b), I cannot perceive the surface of Y dorsally. But because of (c), I need to perceive the surface of Y dorsally in order to see Z in Y. Thus, I cannot see Z in Y.[14]

My response is to bite the bullet. We do not see anything in a depicted picture: we do not see Z in Y. We do see Z in X, the picture we are looking at, and whose surface we do perceive dorsally. And we do see Y in X. We see a picture (picture Y) in Z and we see the object depicted by this picture Y in X, but we do not see Z in Y. Consider the painting on the wall of the room in Vermeer's *Woman holding a balance*

[14] It is important to note that this problem does not arise if the surface of X is the same as that of Y, as they most often are when we look at reproductions of paintings. I will take it for granted in what follows that the two surfaces are different and distinguishable.

(National Gallery of Art, Washington, DC). It depicts the Last Judgment. My claim is that we do not see the Last Judgment in the picture on the wall of the woman's room. We do see a painting on the wall of the woman's room in the Vermeer painting and we do see the Last Judgment in the Vermeer painting.

If we want to adjudicate between my explanation of seeing pictures in pictures and the one that is supposed to be problematic for my account, according to which we see things in depicted pictures, we need to decide whether we see things in the depicted picture (on the wall in Vermeer's painting) or in the picture we are looking at (the Vermeer painting itself). I will argue that we have good reason to believe that it's the latter.

Consider the perception of pictures in pictures where a picture is depicted from an oblique angle. Our estimation of the spatial relations of the objects depicted by the depicted picture is notoriously wrong. And our estimation of the spatial relations of objects depicted by a picture viewed from the same oblique angle is not (Vishwanath et al. 2005).

Take Canaletto's *Grand Canal looking South-east from Theampo Santo Sophia to the Rialto Bridge* (Staatliche Museen in Berlin). There is a larger building on the right hand side of the Canal. If we look at this picture face to face, even from a very oblique angle, we have no problem identifying what building on the left hand side of the Canal would be opposite this large building. If, however, we look at a good quality high resolution photograph of this painting that was made from an oblique angle, it is extremely difficult to tell, which building on the left hand side of the Canal is the one that would be opposite the large building on the right.[15]

When looking at the photograph of the painting, we do not perceive the surface of the painting dorsally. Thus, my account predicts that we cannot see objects in this painting in the same way as we would if it were in front of us so that we could perceive its surface dorsally. And this is exactly what we experience: the way we see objects in surfaces that we do not perceive dorsally is distorted. The way we see objects in surfaces that we do perceive dorsally is not distorted even if we look at these pictures from an oblique angle.

Thus, we see the (distorted) Grand Canal in the photograph and we see the painting in the photograph. We do not see the (distorted) Grand Canal in the painting. The "pictures of pictures" objection is avoided.

Conclusion: twofoldness revisited

I have argued that it is necessary for picture perception that our ventral vision attributes properties to the depicted object, whereas our dorsal vision attributes properties to the picture surface. And these separate perceptual processes constitute

[15] These observations can be reproduced by anyone with a poster and a digital camera. In experimental conditions, similar results were shown. See Pirenne (1970) and Sedgwick and Nicholls (1993). Koenderink et al. (2004) seem to suggest the opposite, but it is important to note that the experiments in Koenderink et al. (2004) are supposed to demonstrate that we "always see an object depicted in a frontoparallel pose […] as facing us squarely, whatever the angle of view" (p. 526). In contrast, what I am interested in is not whether we experience the depicted scene as "facing us sqarely," but whether our experience of the spatial relation between the depicted objects is distorted. And Koenderink et al. (2004) do not talk about how our experience of the spatial relation between the depicted objects changes.

the two folds of our twofold experience of pictures. What Wollheim calls the "configurational" aspect of our experience, the perception of the surface, is a dorsal phenomenon, whereas what he calls the "recognitional" aspect of our experience, the perception of the depicted object, is a ventral phenomenon. Thus, the account I outlined here is consistent with, and could be thought of as an extension of, Wollheim's account of our experience of pictures.

Wollheim stated that the twofoldness of our experience is a necessary condition for picture perception. This suggestion has often been dismissed and criticized for its obscurity: it is not clear what Wollheim meant by "twofoldness". I argued that there is a straightforward and empirically plausible way of filling in the details of Wollheim's notion of twofoldness that would preserve the spirit of the original notion but would also make it testable and, therefore, falsifiable.

The difference between perceiving an object face to face and perceiving an object in a picture can then be summarized in the following manner. When we see an object face to face, our dorsal and ventral visual subsystems attribute properties to the same object: the perceived object. When we see objects in pictures, in contrast, the dorsal and the ventral visual subsystems attribute properties to different objects. The ventral subsystem attributes properties to the depicted scene whereas the dorsal subsystem attributes properties to the surface of the picture.

Finally, it is worth highlighting the similarities and differences between the account outlined here and Mohan Matthen's treatment of picture perception in his (Matthen 2005). To my knowledge, it was Matthen who first suggested the possibility that the ventral/dorsal distinction may be relevant for understanding what happens if we perceive pictures—and I owe the basic idea of this paper to him. But there are four crucial differences between his account and mine that are important to highlight. First, Matthen claims that the surface *can* be represented by dorsal perception, but he does not make the stronger claim I argued for, namely, that the surface *must* be represented by dorsal perception in order for us to be able to see objects in this surface.[16] Second, as we have seen in "The depicted object is not represented by dorsal perception," what he means by the "localization of an object in one's egocentric space" is different from what I mean by this concept. Third, Matthen's aim in Chapter 13 of Matthen (2005) is to clarify how we localize (or fail to localize) objects in pictures and to explore how this relates to the visual reference in the case of episodic memory—a theme he developed further in Matthen (2010). He is not concerned with giving an account of picture perception—in fact, he is only interested in defending claim (b) above—the claim that depicted objects are not represented dorsally. Fourth, as a result, he draws an explicit parallel between picture perception and visual imagery (Matthen 2005, pp. 313–314): neither involves dorsal perception of the perceived (or visualized) object—a claim we may have independent reasons to resist (Nanay 2010c).

Acknowledgments I am grateful to Dom Lopes, Mohan Matthen, Joshua Johnson, Nola Semczyszyn, and two anonymous referees. I gave a version of this talk at the 45th Cincinnati Philosophy Colloquium in 2008 and received very useful feedback, especially from Greg Currie, Noel Carroll, John Kulvicki, Mark Rollins and Jesse Prinz.

[16] Matthen said that he would be open to such claim though—personal communication.

References

Aglioti, S., DeSouza, J. F. X., & Goodale, M. A. (1995). Size-contrast illusions deceive the eye but not the hand. *Current Biology, 5*, 679–685.

Ballard, D. H. (1996). On the function of visual representation. In Kathleen Akins (Ed.), *Perception* (pp. 111–131). New York: Oxford University Press.

Brewer, B. (2006). Perception and Content. *European Journal of Philosophy, 14*, 165–181.

Brooks, R. A. (1991). Intelligence without representation. *Artificial Intelligence, 47*, 139–159.

Bruno, N. (2001). *When* does action resist visual illusions? *Trends in Cognitive Sciences, 5*, 385–388.

Bruno, N., & Bernardis, P. (2002). Dissociating perception and action in Kanizsa's compression illusion. *Psychonomic Bulletin and Review, 9*, 723–730.

Busey, T. A., Brady, N. P., & Cutting, J. E. (1990). Compensation is unnecessary for the perception of faces in slanted pictures. *Perception & Psychophysics, 48*(1), 1–11.

Campbell, J. (2002). *Reference and consciousness*. Oxford: Oxford University Press.

Campbell, J. (2011). Sensorimotor knowledge and naïve realism. *Philosophy and Phenomenological Research, 76*, 666–673.

Carroll, N. (1995). Towards an ontology of the moving image. In C. A. Freeland & T. E. Wartenberg (Eds.), *Philosophy and film*. New York: Routledge.

Chemero, A. (2009). *Radical embodied cognitive science*. Cambridge: MIT Press.

Clark, K. (1960). *Looking at pictures*. London: John Murray.

Clark, A. (2001). Visual experience and motor action: are the bonds too tight? *Philosophical Review, 110*, 495–519.

Clarke, T. (1965). Seeing surfaces and physical objects. In M. Black (Ed.), *Philosophy in America* (pp. 98–114). Ithaca: Cornell University Press.

Cohen, J., & Meskin, A. (2008). Photographs as evidence. In Walden Scott (Ed.), *Photography and philosophy: essays on the pencil of nature*. New York: Blackwell.

Currie, G. (1991). Photography, painting and perception. *Journal of Aesthetics and Art Criticism, 49*, 23–29.

Currie, G. (1995). *Image and Mind: film, philosophy and cognitive science*. Cambridge: Cambridge University Press.

Cutting, J. E. (1987). Rigidity in cinema seen from the front row, side aisle. *Journal of Experimental Psychology: Human Perception and Performance, 13*, 323–334.

Daprati, E., & Gentilucci, M. (1997). Grasping an illusion. *Neuropsychologia, 35*, 1577–1582.

Dehaene, S., Naccache, L., Le Clec'H, G., Koechlin, E., Mueller, M., Dehaene-Lambertz, G., et al. (1998). Imaging unconscious semantic priming. *Nature, 395*, 597–600.

Eimer, M., & Schlaghecken, F. (2003). Response facilitation and inhibition in subliminal priming. *Biological Psychology, 64*, 7–26.

Evans, G. (1982). *The varieties of reference*. Oxford: Oxford University Press.

Evans, G. (1985). Molyneux's question. In G. Evans (Ed.), *Collected papers* (pp. 364–399). Oxford: Clarendon.

Feagin, S. L. (1998). Presentation and representation. *Journal of Aesthetics and Art Criticism, 56*, 234–240.

Franz, V. H., Bülthoff, H. H., & Fahle, M. (2003). Grasp effects of the Ebbinghaus illusion: obstacle avoidance is not the explanation. *Experiental Brain Research, 149*, 470–477.

Gentilucci, M., Cheiffe, S., Daprati, E., Saetti, M. C., & Toni, I. (1996). Visual illusion and action. *Neuropsychologia, 34*, 369–376.

Gillam, B. (1998). Illusions at century's end. In J. Hochberg (Ed.), *Perception and cognition at century's end* (pp. 95–136). San Diego: Academic.

Gombrich, E. (1960). *Art and illusion*. New York: Pantheon.

Goodale, M. A., & Humphrey, G. K. (1998). The objects of action and perception. *Cognition, 67*, 181–207.

Goodale, M. A., & Milner, A. D. (2004). *Sights unseen*. Oxford: Oxford University Press.

Goodman, N. (1968). *Languages of art*. Indianapolis: Bobbs Merrill.

Greenwald, A. G., Sean, C. D., & Richard, L. A. (1996). Three cognitive markers of unconscious semantic activation. *Science, 273*, 1699–1702.

Hagen, M. A., Glick, R., & Morse, B. (1978). Role of two-dimensional surface characteristics in pictorial depth perception. *Perception and Motor Skills, 46*, 875–881.

Halloran, T. O. (1989). Picture perception is array specific: viewing angle versus apparent orientation. *Perception & Psychophysics, 45*, 467–482.

Heller, M. A., Brackett, D. D., Wilson, K., Yoneyama, K., Boyer, A., & Steffen, H. (2002). The haptic Müller-Lyer illusion in sighted and blind people. *Perception, 31*, 1263–1274.

Heller, M. A., McCarthy, M., Schultz, J., Greene, J., Shanley, M., Clark, A., et al. (2005). The influence of exploration mode, orientation, and configuration on the haptic Müller-Lyer illusion. *Perception, 34*, 1475–1500.

Hopkins, R. (1998). *Picture, image and experience. A philosophical inquiry*. Cambridge: Cambridge University Press.

Hopkins, R. (2002). Spectator in the picture. In R. Van Gerwen (Ed.), *Richard Wollheim and the art of painting* (pp. 215–231). Cambridge: Cambridge University Press.

Hopkins, R. (2010). Inflected pictorial experience: its treatment and significance. In C. Abell & K. Bantilaki (Eds.), *Philosophical perspectives on depiction*. Oxford: Oxford University Press. in press.

Hutto, D., & Myin, E. (2011). *Radicalizing enactivism*. Cambridge: MIT Press.

Jacob, P., & Jeannerod, M. (2003). *Ways of seeing. The scope and limits of visual cognition*. Oxford: Oxford University Press.

Jeannerod, M. (1997). *The cognitive neuroscience of action*. Oxford: Blackwell.

Jeannerod, M., Decety, J., & Michel, F. (1994). Impairment of grasping movements following a bilateral posterior parietal lesion. *Neuropsychologia, 32*, 369–380.

Koenderink, J., van Doorn, A., Kappers, A., & Todd, J. (2004). Pointing out of the picture. *Perception, 33*, 513–530.

Króliczak, G., Heard, P., Goodale, M. A., & Gregory, R. L. (2006). Dissociaition of perception and action unmasked by the hollow-face illusion. *Brain Research, 1080*, 9–16.

Kulvicki, J. (2006). *On images: their structure and content*. Oxford: Oxford University Press.

Levinson, J. (1998). Wollheim on pictorial representation. *Journal of Aesthetics and Art Criticism, 56*, 227–233.

Lopes, D. M. (1996). *Understanding pictures*. Oxford: Oxford University Press.

Lopes, D. M. (2005). *Sight and sensibility*. Oxford: Oxford University Press.

Martin, M. G. F. (2006). On being alienated. In T. Gendler & J. Hawthorne (Eds.), *Perceptual experience* (pp. 354–410). Oxford: Oxford University Press.

Matthen, M. (2005). *Seeing, doing and knowing: a philosophical theory of sense perception*. Oxford: Oxford University Press.

Matthen, M. (2010). Is memory preservation? *Philosophical Studies, 148*, 3–14.

Maynard, P. (1994). Seeing double. *Journal of Aesthetics and Art Criticism, 52*, 155–167.

Maynard, P. (1996). Perspective's places. *Journal of Aesthetics and Art Criticism, 54*, 23–40.

Maynard, P. (forthcoming). What s so funny? Artefacts for imagining.

Milner, A. D., & Goodale, M. A. (1995). *The visual brain in action*. Oxford: Oxford University Press.

Nanay, B. (2004). Taking twofoldness seriously. Walton on imagination and depiction. *Journal of Aesthetics and Art Criticism, 62*, 285–289.

Nanay, B. (2005). Is twofoldness necessary for representational seeing? *British Journal of Aesthetics, 45*, 263–272.

Nanay, B. (2010a). Inflected and uninflected perception of pictures. In C. Abell & K. Bantilaki (Eds.), *Philosophical perspectives on depiction*. Oxford: Oxford University Press.

Nanay, B. (2010b). Transparency and sensorimotor contingencies: do we see through photographs? *Pacific Philosophical Quarterly, 91*, 463–480.

Nanay, B. (2010c). Perception and imagination: amodal perception as mental imagery. *Philosophical Studies, 150*, 239–254.

Nanay, B. (2011a). Do we perceive apples as edible? *Pacific Philosophical Quarterly 92*, 305–322.

Nanay, B. (2011b). Do we sense modalities with our sense modalities? *Ratio 24*, 299–310.

Nanay, B. (forthcoming a). Action-oriented perception. *European Journal of Philosophy*.

Nanay, B. (forthcoming b). *Between Perception and Action*. Oxford: Oxford University Press.

Newall, M. (2003). A restriction of pictures and some consequences for a theory of depiction. *Journal of Aesthetics and Art Criticism, 61*, 381–394.

Noë, A. (2002). On what we see. *Pacific Philosophical Quarterly, 83*, 57–80.

Noë, A. (2003). Causation and perception: the puzzle unravelled. *Analysis, 63*, 93–100.

Noë, A. (2004). *Action in perception*. Cambridge: The MIT Press.

O'Regan, K. (1992). Solving the "real" mysteries of visual perception: the world as an outside memory. *Canadian Journal of Psychology, 46*, 461–488.

Peacocke, C. (1987). Depiction. *Philosophical Review, 96*, 383–410.

Pirenne, M. H. (1970). *Optics, painting, and photography*. Cambridge: Cambridge University Press.

Podro, M. (1991). Depiction and the golden calf. In N. Bryson, M. A. Holly, & K. Moxey (Eds.), *Visual theory* (p. 163). New York: Harper Collins.

Podro, M. (1998). *Depiction*. Cambridge: Harvard University Press.

Poincaré, H. (1905/1958). *The value of science*. New York: Dover.

Polanyi, M. (1970). What is a painting? *British Journal of Aesthetics, 10*, 225–236.
Prinz, J. (2010). How do perceptual states become conscious? In B. Nanay (Ed.), *Perceiving the world: new essays on philosophy of perception*. New York: Oxford University Press.
Rizzolatti, G., & Matelli, M. (2003). Two different streams form the dorsal visual system: anatomy and functions. *Experimental Brain Research, 153*, 146–157.
Rogers, S. (1995). Perceiving pictorial space. In W. Epstein & S. Rogers (Eds.), *Perception of space and motion* (pp. 119–163). San Diego: Academic.
Schier, F. (1986). *Deeper into pictures*. Cambridge: Cambridge University Press.
Schier, F. (1993). Van Gogh's boots: the claims of representation. In D. Knowles & J. Skorupski (Eds.), *Virtue and Taste*. Oxford: Blackwell.
Sedgwick, H. A., & Nicholls, A. L. (1993). Cross talk between the picture surface and the pictorial scene: effects on perceived shape. *Perception, 22*(suppl), 109.
Siegel, S. (2006). Which properties are represented in perception? In T. Gendler & J. Hawthorne (Eds.), *Perceptual experience* (pp. 481–503). Oxford: Oxford University Press.
Strahan, E. J., Spencerand, S. J., & Zanna, M. P. (2002). Subliminal priming and persuasion: striking while the iron is hot. *Journal of Experimental Social Psychology, 38*, 556–568.
Suzuki, K., & Arashida, R. (1992). Geometrical haptic illusions revisited: haptic illusions compared with visual illusions. *Perception & Psychophysics, 52*, 329–335.
Topper, D. (2000). On anamorphosis: setting some things straight. *Leonardo, 33*, 115–124.
Travis, C. (2004). The silence of the senses. *Mind, 113*, 57–94.
Turnbull, O. H., Driver, J., & McCarthy, R. A. (2004). 2D but not 3D: pictorial depth deficits in a case of visual agnosia. *Cortex, 40*, 723–738.
Vishwanath, D., Girshick, A. R., & Banks, M. S. (2005). Why pictures look right when viewed from the wrong place. *Nature Neuroscience, 8*, 1401–1410.
Walton, K. (1984). Transparent pictures. On the nature of photographic realism. *Critical Inquiry, 11*, 246–277.
Walton, K. L. (1990). *Mimesis and make-believe. On the foundations of the representational arts*. Cambridge: Harvard University Press.
Walton, K. L. (1991). Reply to Reviewers. *Philosophy and Phenomenological Research 51*, 423–427.
Walton, K. (1997). On pictures and photographs. Objections answered. In R. Allen & M. Smith (Eds.), *Film theory and philosophy* (pp. 60–75). Oxford: Oxford University Press.
Walton, K. L. (2002). Depiction, perception, and imagination. *The Journal of Aesthetics and Art Criticism, 60*, 27–35.
Warburton, N. (1988). Seeing through 'seeing through photographs'. *Ratio (new series), 1*, 64–74.
Westwood, D. A., Danckert, J., Servos, P., & Goodale, M. (2002). Grasping two-dimensional images and three-dimensional objects in visual-form agnosia. *Experimental Brain Research, 144*, 1432–1106.
Wollheim, R. (1980). Seeing-as, seeing-in, and pictorial representation. In *Art and its object* (2nd ed., pp. 205–226). Cambridge: Cambridge University Press.
Wollheim, R. (1987). *Painting as an art*. Princeton University Press: Princeton.
Wollheim, R. (1998). On pictorial representation. *Journal of Aesthetics and Art Criticism, 56*, 217–226.