Journal of the American Philosophical Association (2017) 372–393 © American Philosophical Association
DOI: 10.1017/apa.2017.28

Are Pictures Peculiar Objects of Perception?

ABSTRACT: *Are face-to-face perception and picture perception different perceptual phenomena? The question is controversial. On the one hand, philosophers have offered several solid arguments showing that, despite some resemblances, they are quite different perceptual phenomena and that pictures are special objects of perception. On the other hand, neuroscientists routinely use pictures in experimental settings as substitutes for normal objects, and this practice is successful in explaining how the human visual system works. But this seems to imply that face-to-face perception and picture perception are very similar, if not actually the same. How can we decide between these two opposite intuitions? Here I offer a regimentation of the notion of picture perception that can reconcile these two apparently conflicting ideas about pictures. It follows that philosophers and neuroscientists can maintain their respective stances without any theoretical conflict.*

KEYWORDS: picture perception, neuroscience, experimental setting, vision, visual cortex

Introduction

You are at the zoo, looking at a wonderful cheetah that is approaching you. You decide to immortalize the magnificent animal. Thus, you pick up your Polaroid camera and take a picture. The camera prints the picture of the cheetah, which is then extracted from the bottom of the device. You want to check whether the photo shot is satisfactory and immediately lift the picture up to the left of your visual field to compare the depicted cheetah with the real cat, which occupies the right portion of your visual field, in the external environment, behind the cage. On the left you have the depicted cheetah; on the right you have the real one.

This situation recalls an everyday life practice. However, if you are a philosopher and you are interested in perception, especially in picture perception, you know that

This work was supported by the 'Fondazione Franco e Marilisa Caligara per l'Alta Formazione Interdisciplinare'. I have several special thanks to offer. The first goes to Bence Nanay for his numerous crucial suggestions concerning this project. The second goes to two anonymous referees, whose comments have allowed me to improve the first version of this paper. Very special thanks go to these scholars who enthusiastically discussed with me the topics tackled in this paper: Alberto Voltolini, Andrea Borghini, Silvano Zipoli Caiani, Chiara Brozzo, Eris Chinellato, Anna Maria Borghi, Giorgia Committeri, Dan Cavedon-Taylor, Alva Noë, Clotilde Calabi, Alfredo Paternoster, Vittorio Gallese, Fausto Caruana, Corrado Sinigaglia, Neil Van Leuween, Mario Alai, and Pierre Jacob. Finally, special thanks go to the students of philosophy in Urbino, who attended my lessons in 'Philosophy of Mind and Cognitive Science' and offered several points on this topic.

https://doi.org/10.1017/apa.2017.28 Published online by Cambridge University Press

this very usual situation denotes something really puzzling: what is the relation between the perceptual state we are in when we fix the depicted cat in the left portion of our visual field and the one we are in when we shift the direction of our gaze toward the right portion of our visual field, where we can see the real cat? In other words:

> QP1 What is the difference between perceiving an object in a picture and perceiving a real object (see Nanay 2011, 2017)?

An answer to QP1 presupposes an answer to a more primitive question:

> QP2 Are face-to-face perception and picture perception different?

Only a positive answer to QP2 leads to a possible speculation concerning QP1. These questions denote that pictures might be considered as special—or strange (Gregory 2012)—objects of perception. They merit a serious interest not only as objects of aesthetics, but also from a perceptual point of view (Lopes 2005; Hopkins 2010, 2012; Nanay 2016; Briscoe 2016). As Jacob and Jeannerod state:

> Visual experience of the painting raises a number of fascinating issues for cognitive science . . . on the assumption that one is indeed perceiving a visual array consisting of nine objects with their shapes, contours, orientations, textures, colors and intricate spatial relationships. But, of course, none of this is literally true. Your visual system with the rest of your brain is playing a trick on you: there is no apple, no chestnut, no silvery beaker, no brown dish, no spoon, no table, no wall. All there is is a canvas with two-dimensional shapes and patches of colors drawn on it. How on earth does one see three apples, two chestnuts, a silvery beaker, a brown dish and a spoon when there are no such things to be seen? Or are there after all? This is the puzzle of visual art. . . . Human beings . . . can be fooled by *trompe-l'œil* representations, but they have the joint ability to see what is in a picture and to see a picture as a picture. . . . The beginning of an answer to the puzzle of visual art lies in the idea that in seeing what any normal person takes himself or herself to be seeing in . . . one is 'playing a game of make-believe'. One pretends that one is being presented with . . . apples, chestnuts, a dish, a spoon and a beaker, while one knows that there are no such things. One knows that instead there is a representation of these things. (Jacob and Jeannerod 2003: x–xi)

The current theoretical stance about the status of pictures, in relation to QP2, shows a very controversial situation. Indeed, philosophers and cognitive neuroscientists have strongly different background assumptions toward the pictorial world (Nanay 2015). On the one hand, philosophers have suggested that despite the effective resemblances, picture perception and face-to-face perception are quite different

https://doi.org/10.1017/apa.2017.28 Published online by Cambridge University Press

perceptual phenomena for several crucial reasons (Lopes 2005; Wollheim 1998; Nanay 2011, 2015, 2017; Matthen 2005; Ferretti 2016a, 2016c, 2017a). On the other hand, on a daily basis, the neuroscientific practice makes use of pictures in experimental settings as substitutes of normal objects. Arguably, what lies beneath this practice is a particular, theory-laden, lurking philosophical stance concerning a possible negative answer to QP2. This stance ignores the philosophical reasons related to the positive answer to QP2. How can we decide which is the right view? Here are the reasons that the neuroscientist and the philosopher might offer each other in order to justify their respective practices.

The philosopher might warn the neuroscientist that given the differences suggested in the philosophical literature, neuroscientists are not justified in pursuing the practice of using pictures without paying attention to the philosophical reasons that support a positive answer to QP2 because this might lead to both theoretical and empirical problems. But the neuroscientist might reply to the philosopher by mentioning the following fact: there is compelling evidence that both looking at objects in a picture and looking at normal objects foster a very similar, if not the same, behavior of the cortical visual system. I will get back to this in (section 3). This fact is sufficient to justify the neuroscientific practice. Let us resume these two points with two different theses about picture perception:

> NT (*The Neuroscientific Thesis*): Pictures and normal objects set us in the same perceptual state.
>
> PT (*The Philosophical Thesis*): Pictures and normal objects do not set us in the same perceptual state though there are several resemblances between face-to-face and picture perception.

Empirically informed philosophical accounts of picture perception that recognize and use the evidence reported by neuroscience about the several resemblances of face-to-face and picture perception (Nanay 2011, 2015; Ferretti 2016a, 2016c, 2017a) suggested that, despite such important resemblances, face-to-face and picture perception are different.

Interestingly, even some neuroscientists recognized that the neuroscientific practice is rather odd from a theoretical point of view. Indeed, the vast majority of fMRI evidence aimed at finding the neural correlates of human visual processing is cashed out from experimental settings involving 2-D objects (Snow et al. 2011). Here is the curious thing. In everyday life we usually see and interact with real objects we visually represent as being present in our action space. The human visual system has evolved and developed its functions with respect to the presence of real objects. And humans have no problem in distinguishing between pictures and real objects, which notwithstanding the similarities are different things. This perceptual fact does not seem to be taken into account by neuroscientists, and this stance is common in imaging studies used to map the neural correlates of vision. Indeed, there is poor fMRI evidence based on ecological experimental settings in which the stimuli are real objects (Snow et al. 2011). Curiously, behavioral results from vision science suggest that this stance presents several methodological problems (section 2).

https://doi.org/10.1017/apa.2017.28 Published online by Cambridge University Press

Philosophers have offered very technical and convincing arguments in favor of PT. But the results coming from neuroscience—in line with NT—seem to give us reliable evidence that human visual processing perfectly encodes pictures so that they can be considered licit substitutes of normal objects. The challenge is to cash out an explanation that justifies both positions at the same time (a similar reconciliation has been offered in relation to the special case of trompe l'oeil pictures, see Nanay [2015]), while avoiding the apparent paradox stemming from the comparison of NT and PT. How can we offer such an account?

To my knowledge, no previous philosophical attempt has tried to tackle this theoretical issue. Here I reconcile neuroscience and philosophy in relation to this important point. I will suggest that, if appropriately regimented, these two theses are not really in conflict. Thus (a) neuroscientists are justified in their everyday practice with pictures; (b) philosophers are justified in arguing for the peculiar status of pictures as particular objects of perception.

Here is the plan for the paper. First, I specify that pictures are composed of two things: a picture's surface or vehicle and a depicted scene or object. Thus, while in face-to-face perception our visual system visually represents only one thing, namely, the object, during picture perception it represents two things: the picture's surface and the depicted object (section 1). This specification allows a specific regimentation of the two theses to be offered. Thanks to this regimentation, I will show that we can embrace both PT (section 2) and NT (section 3) without any contradiction (section 4).

1. Pictures, Surfaces, and Depicted Objects

In this paper I want to defend the compatibility between NT and PT, which at first sight seem to be in conflict. Here, I offer a careful analysis and a regimentation of these two theses able to reveal that they are not in opposition. This will allow me to suggest that neuroscientists are justified in their everyday practice with pictures and philosophers are justified in arguing for the peculiar status of pictures as particular objects of perception.

The regimentation I propose is based on a background specification that turns out to be crucial. The fact that when perceiving pictures we perceive their two components, namely, a picture's surface or vehicle and a depicted scene or object, leads us to a very important distinction. It is one thing to talk about picture perception, which refers to both the perception of the surface and of the depicted object. It is another thing to talk about the perception of the depicted object, which is just one aspect of the phenomenon of picture perception. The other aspect is the perception of the surface. Thus, it seems that picture perception is different from face-to-face perception, in which our visual system represents only one thing: the object. This 'twofold' aspect of picture perception is uncontroversial in the literature (Hopkins 2012; Lopes 1996, 2005; Kulvicki 2006; Wollheim 1998; Levinson 1998; Nanay 2011, 2015, 2017; Matthen 2005; Abell 2009; Abell and Bantinaki 2010; Briscoe 2016; Hecht et al. 2003, see section 2). Naturally, I am not concerned with imagination and nonperceptual representations in picture perception here

https://doi.org/10.1017/apa.2017.28 Published online by Cambridge University Press

(Nanay 2017). Of course, during pictorial experience, these two components are not sharply separated from a perceptual point of view. However, much of the philosophical reflection that endorses the perceptual fusion of these two aspects investigates picture perception by analyzing the way in which different portions of our visual system are respectively attuned to the depicted object and to the surface (Nanay 2011: 464, 2015, 2017: sect. 2; Lopes 2005; Hopkins 2012; Ferretti 2016a, 2016c, 2017a). This is because 'any account of seeing-in must be able to tell how this experience represents the picture's surface and how it represents the depicted object' (Nanay 2010b: 199, cf. Hopkins 2010). Here I endorse this methodology. Accordingly, vision science shows that the presence of a perceptual fusion between the two aspects, namely, the perception of the surface and the one of the depicted object, does not prevent us from examining, with specific experimental techniques and settings, how our visual brain builds separate conscious or unconscious visual representations of the surface and of the depiction and from studying them in isolation (Nanay 2011: 463, footnote 1; 2017: sect. 2; see section 2 below). This specification leads me to regiment NT in a sharper way. It is now possible to distinguish between two versions of this thesis:

> *The Strong Neuroscientific Thesis*: *Pictures* and normal objects set us in the same perceptual state.
>
> *The Weak Neuroscientific Thesis*: *Depicted objects* and normal objects set us in the same perceptual state.

We can endorse the *Weak* NT without endorsing the *Strong* NT. *Depicted* objects and *normal* objects foster in us the same visual response. This does not conflict with the idea that *pictures* and *normal* objects do not set us in the very same perceptual state. Indeed, during the perception of depicted objects our visual system can represent several visual properties it also represents in face-to-face perception. However, a depicted object is embedded in a surface. Therefore, unlike face-to-face perception, in picture perception there are two visual representations, one of the surface and the other one of the depicted object—let us call this the *Double Representation Thesis*. If we take these two representations separately—and a few lines above we saw that this move is licit—we can say that the visual representation of a normal object can be similar, from a visual point of view, either to that of a depicted object or to that of the picture's surface (Nanay 2011; Ferretti 2016a, 2017a), except for at least one crucial difference, described below. However, face-to-face and picture perception are overall different phenomena: the former involves only one visual representation, while the latter involves two representations.

There is a crucial point we have to take into account here. As mentioned above, we can endorse the view that during the perception of depicted objects, our visual system can represent several visual properties we also represent in face-to-face perception. This remains true even if depicted objects cannot foster in us the visual feeling of presence because they are embedded in a visible surface, (Ferretti 2016c; Matthen 2005; Nanay 2015) - I will get back to this in section 2. The visual feeling of presence is related to the fact that we have a perception of the object as a present one that we can interact with in our action space (something possible with

https://doi.org/10.1017/apa.2017.28 Published online by Cambridge University Press

the perception of the surface). Thus, we can take into account this difference by sharpening the *Weak* NT:

> *The Sharpened Weak Neuroscientific Thesis*: *Depicted objects* and normal objects set us in a *very similar* perceptual state. It is *not* the *very same state* because we can attribute the same properties to both, except for *at least one* property that we can attribute only to normal objects (and thus only to surfaces, which are real objects): the property of being present.

It is easy to note that the *Strong* NT conflicts with PT. And the *Weak* NT is not sufficiently specific. But if we select the *Sharpened Weak* NT, we can reconcile neuroscience and philosophy. Indeed, saying that depicted and normal objects set us in a very similar perceptual state and that depicted objects can make our visual system represent several visual properties as normal objects do, can justify the neuroscientific practice. Also, this does not conflict with the *Double Representation Thesis*. We are just saying that the visual representation of the depicted object, which is only one of the two representations performed in picture perception—the other one is related to the surface—share many characteristics with the visual representation of a normal object. I will get back to an analysis of the nature of this similarity in section 3. This does not imply that the twofold perceptual processing involved in picture perception is the same perceptual state as the one we are in during face-to-face perception: face-to-face perception does not depend on the representational simultaneity involved in picture perception (section 1). Finally, this takes into account that only face-to-face perception can foster in us the visual feeling of presence. And this is precisely because in the case of picture perception we always represent the picture's surface—the *Double Representation Thesis*. Crucially, as we shall see (section 2), this representation makes it impossible to represent the depicted object as present. All this is in accordance with PT.

In what follows, I first report evidence in support of PT. Then, I report evidence in support of the *Sharpened Weak* NT. Finally, I suggest that if we select the *Sharpened Weak* NT, philosophy and neuroscience can be reconciled without any problem because this is perfectly compatible with PT.

2. The Philosophical Thesis

Here I suggest that philosophers are justified in arguing for the peculiar status of pictures as particular objects of perception. Thus, we can support PT.

Embracing the *Double Representation Thesis* (Lopes 2005; Nanay 2011, 2015, 2017; Ferretti 2016c, 2017a) leads us to a famous debate about whether we simultaneously visually represent the picture's surface and the depicted object (Wollheim 1980) or whether, conversely, we visually alternate between them (Gombrich 1960). Most of the current effort in the literature on picture perception has been geared toward choosing one of these two ideas (Nanay 2011, 2017; Cavedon-Taylor 2011; Chasid 2014; Hopkins 2012; Kulvicki 2006). The balance

https://doi.org/10.1017/apa.2017.28 Published online by Cambridge University Press

leans on Wollheim's idea that we always simultaneously visually represent both the surface and the depicted object. However, it is specified that—different from Wollheim's initial insight—we usually consciously represent the depicted object, but we may or may not consciously see/attend to the picture's surface and, indeed, most of the time we do not (Nanay 2011: 463–64, 2017). We can take this simultaneity 'not as simultaneous attention, but as simultaneous (conscious or unconscious) representations' (Nanay 2015: 192; see also Nanay 2017). This implies, ipso facto, that we do not need to consciously represent both of them simultaneously (Nanay 2011, 2017; Hopkins 2012). (This is a weak interpretation of Wollheim's idea of 'Twofoldness'; Wollheim originally suggested the simultaneity of conscious representation.) I am using 'conscious vision' and 'conscious visual attention' interchangeably because I do not take part on the bigger debate on their relation when talking about picture perception (see Nanay 2011, 2017). The *Double Representation Thesis* follows this weaker interpretation of simultaneity according to which, most of the time in picture perception, we consciously visually represent the depiction and, at least unconsciously, visually represent the surface.

Several arguments support this notion of simultaneity. First, it is possible to have unconscious perception of either the surface or of the depicted object (Nanay 2011: footnote 1). If it were necessary to have conscious perception of both, this would rule out evidence about the possibility of unconscious picture perception; conscious perception of both is not necessary (Nanay 2011: 463; 2017). This is also confirmed by the fact that, most of the time, in our everyday experience of pictures, we tend consciously to focus on the depicted object while ignoring, from a conscious/attentional point of view, the surface. Nonetheless, we always, at least unconsciously, visually represent the surface (Nanay 2011: 463; 2015, 2017; Lopes 2005; Levinson 1998). Finally, the picture's surface is visually represented, but not necessarily consciously because perceiving pictures from an oblique angle does not lead to distortions. This phenomenon is explained by invoking the influence of the unconscious perception of the surface/vehicle on the perception of the pictorial content (Nanay 2011: 471; 2017: 5.1; Matthen 2005: 315–317; Hopkins 2012: 657; for further crucial arguments against the idea of simultaneous conscious representation of both the surface and the depiction see Hopkins [2012] and Nanay [2017]).

There is, however, another crucial argument showing that we always visually represent, at least unconsciously, the picture's surface: the representation of the surface is what prevents us from attributing the visual feeling of presence to the depicted object (Vishwanath 2014; Ferretti 2016c);- there are particular visual features—glossiness, shininess, plasticity, and so on— that we need to detect in order to reach the feeling of presence (see Vishwanath 2014; Ferretti 2016c). Indeed, when we cannot represent the surface, depicted objects look as present as normal objects do. In everyday life, however, depicted objects do not look as present as normal objects do. They look depicted, which is what they really are. Therefore, we are representing the surface.

Vision science has clearly shown this perceptual fact. For example, in an experiment subjects are made to look at a depicted scene monocularly through an oval aperture. The aperture is located in front of the eye such that it occludes the

https://doi.org/10.1017/apa.2017.28 Published online by Cambridge University Press

rectangular boundary of the image (the frame and its boundaries of the picture's surface). In this case, the depicted object is perceived as a normal present object occupying a portion of the space in front of the subject she/he can interact with (Vishwanath and Hibbard 2013: 1674; Vishwanath 2014: 153). This result is also drawn from both psychophysical and phenomenological evidence (Vishwanath 2014: 174; 2011: 224, 225; Vishwanath and Hibbard 2010, 2013; Ferretti 2016c: 2.4; Barry 2009).

Therefore, when the surface is not visible—for example, due to the invisibility of the contours—even depicted objects can look present to the observer (Vishwanath 2014; Ferretti 2016c). In the words of Vishwanath and Hibbard:

> When a picture is viewed normally with both eyes, the picture's surface is visible because of cues such as binocular disparity and the visible frame of the picture. . . . Distance cues such as binocular convergence, vertical disparity, and the accommodative state of the lens specify the distance of this visible picture surface . . . rather than the pictorial contents. . . . There are no known optical cues that specify the distance of pictorial objects from the observer. Therefore, under binocular viewing of pictures, although 3-D object shapes can be clearly perceived, their scale and absolute depth should remain optically unspecified. Monocular aperture viewing removes the main cues that specify the presence of the picture's surface (binocular disparity and the visible frame), as well as binocular cues specifying its distance (convergence and vertical disparity). However, subsidiary distance cues, such as the accommodation state of the lens, are still present. In the absence of visible picture's surfaces, it is plausible that the brain attributes the accommodation response to the pictorial objects, and assigns any associated distance information to them, allowing absolute depth values to be derived. . . . It is likely that viewing conditions that yield only a partial reduction in surface visibility (removing either binocular disparity or visible frame) will result in a greater variability in the assignment of distance information to the pictorial objects and a correspondingly reduced sense of stereopsis. (Vishwanath and Hibbard 2013: 1682–1683; see also Vishwanath 2014: 159–160)

The reader should note that stereopsis is the peculiar visual process that allows us to perceive an object as vividly present in our action space (Ferretti 2016c). It depends on the possibility of representing absolute egocentric depth values (Vishwanath 2014). In face-to-face object recognition, unlike in picture perception, monocular and binocular depth cues, retinal and expected size, and viewing distance are consistent. Additional shape information is also available from stereopsis, which allows us to represent vergence and disparity correctly (Snow et al. 2011).

Furthermore, egocentric absolute depth concerns the fact that the 'observer has knowledge of the depth relations scaled in some meaningful way to the actions of the observer' (Vishwanath 2011: 222; see also p. 206). Ascribing absolute

https://doi.org/10.1017/apa.2017.28 Published online by Cambridge University Press

egocentric depth values to the objects we face in our environment is what leads us to represent visually an object as present in our action space and as offering reliable possibility for motor interaction (Ferretti 2016c). When we cannot visually represent the surface, absolute egocentric depth values are assigned to the depicted objects. Thus, they look present.

This evidence strongly suggests that we do not visually represent depicted objects as present because the surface's visibility plays a crucial role in avoiding the induction of the visual feeling of their presence: the visual representation involved in the recognition of the vehicle/surface as present hinders the possibility of having the visual impression of the presence of the depicted object (Vishwanath 2014: 164; Ferretti 2016c: 2.2). Thus, when we cannot visually represent the former, the latter looks present. But our usual binocular picture perception can properly visually assign the related distances to both the surface and the depiction. Thus, we visually represent the surface, and the depicted object is represented as such, that is, as a non-present pictorial object. This evidence explains why depicted objects do not look present, in line with the *Double Representation Thesis*.

Of course, there are genuine situations of picture perception, built without specific experimental settings, in which we cannot visually represent the surface, and thus, what effectively is a depicted object looks like a present one. One common example is the special case of trompe l'oeil pictures. Trompe l'oeil pictures are built in such a way that (even in binocular conditions and without apertures to look through) our visual system cannot properly detect the picture's surface. Thus, it is fooled by the illusory presence of very enhanced visual depth cues (see Ferretti 2016c) wrongly attributed to the depicted object.

I do not need to focus deeply on this case here since this paper is about normal pictures. But this evidence is crucial because in line with the *Double Representation Thesis* it suggests that, during picture perception, we *do* see—or visually represent—both the surface and the depicted object. Following the evidence from vision science, if we were only representing the depicted object but not the surface, the former would look present, like a real object. Accordingly, in the case of trompe l'oeil perception, we also do not have an unconscious representation of the surface. If there were, at least an unconscious representation of the surface, we would not have the feeling of presence of the depicted object.

One might argue that with trompe l'oeil pictures we do perceive the surface unconsciously, and this might explain the difference between normal perception and perception of trompe l'oeil pictures (Nanay 2010b, 2011: footnote 11, 2015, 2017: 5.1; Ferretti 2016c: sec. 3). However, if this were the case, we would not be able to distinguish the trompe l'oeil case from what is taken to be the most usual case of picture perception, according to the literature, in which, following the notion of simultaneity embraced by the *Double Representation Thesis*, we consciously see the depicted object and perceptually realize it is a depicted object. At the same time, we do not consciously visually represent the surface because, most of the time, our conscious vision is only focused on the depiction (Lopes 2005; Levinson 1998; Nanay 2011, 2017). After all, both of these situations involve an unconscious visual representation of the picture's surface. However, crucially, we know that only one of them, namely, the trompe l'oeil one, fosters in us the feeling of

https://doi.org/10.1017/apa.2017.28 Published online by Cambridge University Press

presence. Therefore, invoking the unconscious representation of the surface cannot be the right explanation for trompe l'oeil perception. The only way to explain the difference is by saying that even if only for a moment our visual system cannot 'find' the picture's surface and also cannot do so at the unconscious level: not only are we not consciously visually representing the surface, as in the case of usual picture perception mentioned above, but our visual system cannot even build an unconscious representation of a surface (Ferretti 2016c; Vishwanath 2014; Nanay 2017): [1] it cannot track that surface's presence.

This point is very useful for understanding the importance of the presence, in usual picture perception, of an *at least unconscious* visual representation of the surface, which accompanies our conscious visual representation of the depicted object. The presence of this unconscious representation allows us to focus on/attend to the depicted object consciously without visually representing it as present. However, this representation also guarantees that it is possible to shift our conscious visual attention to the picture's surface, which is, then, consciously visually represented (Nanay 2017; i.e., the unconscious visual representation of the surface can sometimes be consciously accessed; I have little to say about this case here, which is usually called *inflection*, see Voltolini 2013, Lopes 2005, Nanay 2010b). The reader should note that the notion of shift might recall Gombrich's notion of shift. However, his idea is about visual representation, while here shifting concerns conscious vision. (This is not possible, or at most only momentarily, with trompe l'oeil because we have no visual representation, not even an unconscious one, of the surface of the trompe l'oeil picture.

Explaining pictorial unconscious representations is crucial because 'philosophers are trying to explain the experience of seeing pictures whereas psychologists are trying to explain the (conscious or unconscious) mechanism of picture perception. In fact, unconscious picture perception is oddly missing from philosophical discussions of picture perception' (Nanay 2015: 183).

Of course, another difference between perceiving face-to-face and perceiving depicted objects is that, in the latter case we can change spatial position in relation to the picture, but this spatial shift does not reflect any change in our perception of the depicted object (Nanay 2011: 471; 2017: 5.1; Matthen 2005: 315–317; Hopkins 2012). Following the title of the paper by Vishwanath and colleagues (2005) 'pictures look right when viewed from the wrong place'. This suggests, as mentioned above, that we represent (even unconsciously) the surface. Indeed, there is no sensorimotor understanding (Noë 2004) with depicted objects because there is no movement for which the observer's view of the depiction changes as she/he

1 Another argument about the difference between normal and trompe l'oeil pictures is that with *the latter* dorsal vision-for-action is active, while with the former it is not (Nanay 2015). However, some have suggested that dorsal vision-for-action (which serves the representation of action properties and motor commands) is always active, even with objects depicted normally (Ferretti 2016a; see also point (b) in section 3)—and this point can coexist (Ferretti 2016a: sect. 4.2; 2016c: sect. 3; 2017) with the dorsal/ventral account of picture perception (Nanay 2011, 2015). Thus, the difference is that with trompe l'oeil pictures our ventral visual recognition cannot visually track the presence of any surface, not even at the unconscious level, and this fosters in us the impression, even at the conscious level, that action performance with the depicted object is possible (Ferretti 2016c; for further details see (b) in section 3).

https://doi.org/10.1017/apa.2017.28 Published online by Cambridge University Press

moves (Nanay 2010a). I will not analyze this evidence because neuroscientists usually use pictures in static experimental settings for discrimination tasks, in which neither the subjects nor the objects presented on the monitor can move (Ferretti 2016a). I do not deny that this is a crucial feature of perceptual presence in face-to-face seeing, especially concerning perceptual learning of sensorimotor contingencies (Noë 2004). However, after learning has properly developed, presence can be perceived even in static cases of face-to-face perception. And the feeling of presence remains a crucial perceptual feature that is missing in static experimental settings with depicted objects but, as we will see, with no crucial problems for several kinds of perceptual attributions.

My account is not in conflict with those who argue that picture perception and face-to-face perception are 'experiences of the same psychological kind' (Briscoe 2016) because even they recognize the impossibility of assigning absolute egocentric depth and presence to pictorial objects. It has also been suggested that picture perception is not special because it resembles face-to-face perception: 'a picture depicts an object by generating in those who view the picture a visual experience of that object' (Chasid 2014: 471). This view follows Gombrich's stance opposing Wollheim's idea of simultaneous experience of both the surface and the depiction. However, first, I am not committed to this notion of twofoldness: the surface can be unconsciously represented. Second, these authors do not deny that depicted objects cannot solicit the feeling of presence.

Summing up, unlike normal objects, depicted objects are not represented as present. This is because the representation of the surface prevents us from representing the depicted object as present. However, without at least an unconscious representation of the surface, even depicted objects would always look as present as normal objects do. But this would imply that there is no possibility of experiencing usual pictorial perception, but only naturalistic, *trompe l'oeil*-like perception.

This makes pictures special objects of perception and supports the PT. We now have to examine the arguments in favor of the *Sharpened Weak* NT.

3. The Sharpened Weak Neuroscientific Thesis

Here I defend the idea that neuroscientists are justified in their everyday practice with pictures. Thus, it is possible to endorse the *Sharpened Weak* NT. This idea is supported by empirical evidence that depicted objects do, indeed, induce a very similar behavior of our visual system to that induced in face-to-face perception by normal objects (Nanay 2011, 2015; Ferretti 2016a, 2016c; 2017a). This is possible without the need for depicted objects to be visually represented as present.

But what exactly does it mean that depicted objects induce a similar behavior of our visual system with respect to the one induced by normal objects? Or, conversely, what does it mean that face-to-face perception and the perception of depicted objects set us in a very similar perceptual state?

We perceive a lot of things in the visual scene. Following the usual jargon in the philosophical literature, when we see, for example, a shape, we are in a perceptual

https://doi.org/10.1017/apa.2017.28 Published online by Cambridge University Press

state that is responsible for such a shape perception. This allows us to individuate a visual state on the basis of what it represents (colors, shapes, texture, etc.). Usually, representation is described in terms of property attribution: a visual state represents a certain shape in the sense that through this representation we attribute a given shape property (e.g., the property of being cubic) to a given object. We can also take into account the results from vision science and say that a given visual state depends on the activity of a specific portion of the visual cortex. I am endorsing here a common view about perception, which can be used also for picture perception and according to which perception is described in terms of representations that attribute properties to normal objects (Siegel 2007, 2010; Nanay 2013) and depicted objects (Nanay 2011; Ferretti 2016c, 2017a).

That said, there is compelling evidence that our visual system sets us in a *very similar* perceptual state during both face-to-face perception and the perception of depicted objects. In other words, it attributes *very similar* properties (color properties, shape properties, etc.) to both normal and depicted objects. *The Sharpened Weak* NT perfectly takes into account this perceptual fact. Why *very similar* and not the *very same* perceptual state? Or, to put it in other words, why *very similar* but not the *very same* property attribution?

On the one hand, overall, the perception of depicted objects and face-to-face perception are very similar: we attribute several similar properties to both normal and depicted objects. But as we saw in section 2, there is at least one property that our visual system cannot attribute to depicted objects but only to normal objects: the property of being present (Ferretti 2016c; Matthen 2005; Nanay 2015).

This distinction underlines the number of properties we attribute in these two perceptual situations. There is a set of properties we equally visually attribute to both normal and depicted objects. And there are properties visually attributed only to normal objects. However, the first set is much richer than the second (Briscoe 2016; Ferretti 2016a, 2016c, 2017a; see also below). And the most evident and crucial different property within this second set is the property of being present[2]—this does not conflict with the idea of pictorial presence that 'one visually perceives something present in the realm of the picture when experiencing a picture as a depiction' (Aasen 2015: 188). The possibility of this kind of property attribution distinguishes these two perceptual processes. This is not a problem for the *Sharpened Weak* NT, which perfectly takes into account this perceptual fact.

2 I am not denying that this property may come along with other particular spatial properties that are related to it (section 2). However, unlike the property of being present, some of these spatial properties are not completely absent with depicted objects but only diminished in their visual quality or degree. Consider depth: we can ascribe it to some extent to depictions, but not as deeply as we do with normal objects. Indeed, we can represent *relative* depth among objects in the depicted scene (see (c) below) but not *absolute egocentric* depth, which is related to the observer's action space and is responsible for the feeling of presence (section 3), which is not attributable to depicted objects. This second high visual information is not *sufficiently specified* with depicted objects because they are embedded in a surface (section 2). Normal objects offer different sources of additional shape and spatial information with respect to depicted objects (Briscoe 2016). In line with this point and with the analysis offered in this paper, the use of objects in pictures as substitutes of normal objects without any proper consideration of their differences with respect to volumetric and depth properties has sometimes led neuroscientists to build experimental settings that have been judged by philosophers to be improper with respect to the goal of the experimental test (Schwenkler 2013).

https://doi.org/10.1017/apa.2017.28 Published online by Cambridge University Press

Furthermore, we can address a perceptual hierarchy with respect to both visual representations and property attributions in relation to the distinction between face-to-face perception, picture perception, and the perception of depicted objects. First, only picture perception entails the *Double Representation Thesis*. Second, face-to-face perception and the perception of the depicted objects are, overall, very similar, but not precisely the very same perceptual process; though they attribute very similar properties to their respective objects, only the former but not the latter can attribute the property of being present to its proper perceptual objects. However, and this is an important point, concerning those properties we can attribute in both face-to-face perception and the perception of depicted objects—apart from the feeling of presence—the property attribution is, from a perceptual point of view, the same. This means that, for example, we can equally visually attribute the property of being red to both a depicted and a normal object; this holds for all the properties we can visually attribute to both normal and depicted objects I will list below (I cannot develop this view in full detail here, but see Nanay 2011; Briscoe 2016; Ferretti 2016a, 2107; Hecht et al. 2003). Crucially, this is possible even if depicted objects do not look present.

Summing up, this explanation perfectly captures the idea that we can attribute the same properties to both normal and depicted objects, except for the property of being present. Now we can examine, step by step, the evidence about the similarity between face-to-face perception and the perception of depicted objects, and this evidence shows that we can visually attribute most of the properties we usually attribute to normal objects also to depicted objects and that we can do so even without attributing the property of being present to depicted objects.

First of all, neuroscience teaches us that our visual system is divided, from an anatomo-functional point of view, into two main visual pathways (Milner and Goodale [1995] 2006; Jacob and Jeannerod 2003): one ventral stream for conscious visual recognition in an allocentric frame of reference and one dorsal stream for visually guided action in an egocentric frame of reference. Both streams respond to depicted objects in the same manner as they respond to normal objects. Therefore, most of the perceptual properties usually attributed by each stream to normal objects can also be attributed to depicted objects (Nanay 2011, 2015; Ferretti 2016a, 2016b, 2017a). Here is the list of properties we can visually represent depicted objects as having:

(a) (*Ventral) Recognition of Visual Properties*: When we look at depicted objects, we recognize what they depict. The ventral stream serves conscious or unconscious visual recognition. Thus, the ventral stream is activated during depicted object recognition (Nanay 2011: 467). First, by being activated even during the perception of the depicted object, the ventral stream attributes color properties, space properties, texture, shape, and size properties to the depicted object in the same way it does with normal objects (Grill-Spector and Weiner 2014). Second, cases of cortical damage show that subjects with an impaired ventral stream cannot see things in a picture (Turnbull et al. 2004; Westwood et al. 2002; Nanay 2011: 468). For example, in the case of achromatopsia (the inability to see colors) and prosopagnosia (face blindness) due to specific damage to specific portions of the ventral stream, subjects cannot appreciate depicted colors and faces

https://doi.org/10.1017/apa.2017.28 Published online by Cambridge University Press

(Jacob and Jeannerod 2003: 9, see also pp. 138, 173; for a full account see Nanay 2011).

(b) (*Dorsal) Ascription of Action Properties*: The dorsal stream also represents depicted objects. There are specific dorsal portions that are responsible, during face-to-face perception, for the visuomotor transformation of an object's geometrical attributes into action properties (anterior intraparietal area) and, then, of action properties into motor commands (the most rostral part of the ventral premotor cortex, named F5; Castiello 2005; Sakata et al. 1995; Murata et al. 2000; Raos et al. 2006; Ferretti 2016a, 2016b, 2016c, 2017b). The visuomotor transformation is accompanied by a neurophysiological motor simulation of the motor act related to the action properties (Ferretti 2016a: sect. 4, 2017b). This is a form of covert mental action and is a mechanism of motor preparation for potential action that is simultaneously blocked and can be recruited if the person decides to act. Note that dorsal computations are triggered without the subject wanting or attempting to grasp the object: observation in static conditions directly triggers these dorsal visuomotor responses (Ferretti 2016a: sect. 4). Now, dorsal processing responds not only to real objects but also to depicted objects by attributing to them action properties and representing the related motor commands. Indeed, dorsal processing is triggered regardless of whether an object is real or depicted (Chao and Martin 2000; Proverbio et al. 2011; Buccino et al. 2009; Costantini et al. 2010; Grézes and Decety 2002; Zipoli Caiani 2013; Romero et al. 2014; Pani et al. 2014; Sakata et al. 2003). The evidence for this perceptual fact is well established in both the neuroscientific (ibid.) and the philosophical literature (Zipoli Caiani 2013; Ferretti 2016a, 2016b, 2016c, 2017a) and is compatible (Ferretti 2016a: 3811–12; 2017a: sect. 3.1) with other investigations of the behavior of the visual streams in picture perception (Nanay 2011, 2015).

I am not saying that the *motor behavior of the subject in general* is the same with both real and depicted objects: we do not attempt to grasp objects in a picture. I am saying that *the representational behavior of the dorsal stream* is the same with both of them. Dorsal visuomotor responses are very specific: the same triggered motor response that encodes a specific motor act (e.g., a power grip) in the case of a specific real object (a real handle) is activated in the case of a specific depiction of the very same object (e.g., a depicted handle). The difference in relation to the perception of a depiction and of a real object, is that the result of dorsal computations, related to visuomotor transformation and simulated covert action, remains stored in the motor memory and cannot be used for overt action with depictions. It can be exercised only with real objects when we decide to act (Ferretti 2016a; Zipoli Caiani 2013). Moreover, both the visuomotor transformation and the simulation are automatic unconscious subpersonal processes that are not consciously accessed by the subject (Ferretti 2016a, 2017b).

Thus, saying that our dorsal stream can represent depicted objects as offering action possibilities and, thus, that the representational behavior of dorsal processing in particular, not the motor behavior of the subject in general, is the same with respect to both normal and depicted objects does not mean, of course, that we have the conscious impression that we can act upon those possibilities. It only means that our dorsal brain behaves at the level of automatic subpersonal visuomotor

https://doi.org/10.1017/apa.2017.28 Published online by Cambridge University Press

processing that is not consciously accessible as if the action properties pertaining to the geometrical configuration of the depicted object pertained to a real 3-D object (Ferretti 2016b: 3804). And this is possible because, as I pointed out, dorsal processing is active regardless of whether an object is real or depicted. At the conscious level, ventral processing informs us that we are in front of a depiction, not in front of a present object. And this is because, at the unconscious level, ventral processing tracks the presence of the surface.

This perceptual fact is possible due to the different and peculiar computational characteristics of each stream. Ventral vision is the one involved in the visual selection of targets for reliable action, that is, it is involved in 'action planning'. This is because it is the one involved in 'response selection', a process of high-level cognition that can be conscious or unconscious and that allows us to detect which objects are real and present (that is, nonpictorial), and thus offer reliable interaction and which are not (Westwood et al. 2002; Ferretti 2016a, 2016c). Only ventral vision can perform both pictorial recognition and recognition of presence. As I said, dorsal vision, alone, cannot distinguish whether an object is real or pictorial, and on the basis of the geometrical properties detected, it responds to the depicted objects that are apparently located in the peripersonal space of the observer and, thus, can be represented as apparently reachable (for a complete account see Ferretti 2016a, 2016c).[3] For this reason and because of the fact that due to the magnocellular advantage dorsal computations for motor programming are triggered before ventral processing for recognition occurs (Milner and Goodale [1995] 2006; Barrett and Bar 2009), dorsal visuomotor attribution of action properties and motor simulation occur before ventral response selection and action planning track presence and remain stored in our motor memory for any possible recruitment related to overt action.

Accordingly, only when ventral processing performs response selection and recognizes the object as being present, does it select that object for action and triggers action planning toward it. And only at this point can the thin motor parameters computed by the dorsal stream and stored in our visuomotor quiver be used for overt visuomotor interaction (Milner and Goodale [1995] 2006; Zipoli Caiani and Ferretti 2017; Briscoe and Schwenkler 2015; Briscoe 2009; Goodale and Milner 2004; Chinellato and del Pobil 2016; Ferretti 2016b, 2016c).

On this matter, the literature distinguishes between 'ventral action planning', which establishes whether the object we face is a real one or merely a depiction and, thus, whether it can be selected for action (and the general way to act with respect to our intentions), and 'dorsal motor processing', which computes the specific, thin motor parameters to be used for overt interaction (Dijkerman et al. 2009; Milner and Goodale [1995] 2006: 244; Briscoe and Schwenkler 2015: 1437), triggered independently of whether the object is real (Ferretti 2016a).

3 Also, subjects with dorsal damage fail to evaluate particular geometrical and spatial aspects in the realm of depictions (Nanay 2011: 475; Jeannerod et al. 1994: 370; Jeannerod 1997: 62).

https://doi.org/10.1017/apa.2017.28 Published online by Cambridge University Press

The interplay between these computational characteristics of the streams[4] leads our visuomotor system to respond also to depicted objects without the risk that we attempt to grasp them. Thanks to its computational characteristics, ventral response selection is attuned both to the surface and the depiction (compare (a) above and see below). It recognizes the depicted object as such, namely, as a nonpresent pictorial object. This representation is usually conscious in picture perception. But it also recognizes the surface as a real and present object with which we can interact. This representation is usually unconscious in picture perception (Nanay 2011; Ferretti 2016c). And following what I said above, only in the case of the recognition of presence, given by ventral processing, which in picture perception is related to the surface, is it possible to make use of the motor parameters computed by dorsal processing (Ferretti 2016c). An exception is made for pictorial illusions (e.g., trompe l'oeil pictures, see section 2 and footnote 1) able to deceive our ventral visual recognition that there is no surface. In this case, what effectively is a depiction is perceived by ventral processing as a present object toward which action planning can be triggered and to which motor execution, whose parameters are computed by dorsal processing, can be subsequently directed. All this suggests that, in order not to be fooled, at the conscious level, that the depicted object is present, as in the case of trompe l'oeil, we need *at least* an unconscious representation attuned to the surface of a particular kind: one served by ventral processing and involved in visual recognition of presence, response selection, and action planning (cf. section 2, footnote 1). As explained above, when we can rely on such a representation following the evidence reported in this section, ventral vision allows us to consciously represent the depiction as such, that is, as a pictorial nonpresent object. This is in line with the idea that unconscious recognition of the surface modulates the conscious representation of the depiction (Nanay 2017: sect. 2) and is perfectly compatible with the weaker notion of simultaneity in picture perception endorsed here following the *Double Representation Thesis* (section 1).

All this suggests that the main differences between real and depicted objects are tracked by ventral recognition (Ferretti 2016c). This point is compatible with the claim that both dorsal and ventral vision represent both the depiction and the surface (Nanay 2011; Ferretti 2016a, 2016c, 2017a, see also point (a) above), but only ventral vision represents them, respectively, as a pictorial object and as a real and present one; dorsal vision cannot distinguish between these two different kinds of perceptual objects. But this does not prevent dorsal vision from representing the different motor acts respectively related to the different geometrical arrangements of the depiction (e.g., a depicted handle recalls a power grip) and of the surface (which usually recalls a precision grip) (see Ferretti 2016a, 2016c, 2017a for this point and for the explanation, in line with Nanay's account, of how ventral and dorsal vision are respectively attuned to both the surface and the depiction).

(c) *Ascription of Spatial Characteristics*: We can also represent several spatial features in the pictorial space that resemble those represented during face-to-face

4 Note also that the two streams interact concerning different tasks and in both face-to-face (Chinellato and del Pobil 2016; Briscoe 2009; Briscoe and Schwenkler 2015; Ferretti 2016b) and picture perception (Ferretti 2016c). The account proposed here is perfectly compatible with this idea.

https://doi.org/10.1017/apa.2017.28 Published online by Cambridge University Press

perception. We can perceive 'metric relative depth relations within and among objects, including 3-D surface shape, slant, depth order, and so on, all of which imply an understanding of the ratios of distances or depths among points in the visual field' (Vishwanath 2011: 222). Accordingly, we can perceive orientation, apparent 3-D configuration, layout and volumes (Gonzalez et al. 2008; Snow et al. 2011; Berryhill and Olson 2009; Marotta and Goodale 2001; Cutting 2003; Koenderink and van Doorn 2003; Koenderink 2012; Nelissen et al. 2009; James et al. 2002; Taira et al. 2001). In addition, ventral vision allows us to quasi-egocentrically perceptually represent depicted objects (Briscoe 2009: 447, sect. 6.1–6.4; Grush 2000; Pani et al. 2014; Bruzzo et al. 2008; Vishwanath 2014; Cutting 2003; Hecht et al. 2003): the subject 'non-reflectively assumes the perspective of the camera' (Briscoe 2009: 447). This is in line with the possibility of (apparent) localization of depicted objects in our peripersonal space by the dorsal stream (Ferretti 2016a) and with the fact that absolute egocentric localization is not possible with depictions (see section 2).

(d) *Ascription of Emotional Properties*: The two streams are connected to two different portions of the orbitofrontal cortex, which is involved in the emotional encoding of the visual scene: the dorsal stream is connected to the medial orbitofrontal cortex, while the ventral stream is connected to the lateral orbitofrontal cortex. The former pathway encodes emotional properties salient for visually guided action; the latter encodes emotional properties salient for visual recognition (Barrett and Bar 2009; O'Reilly 2010; Elliott et al. 2000). There is compelling evidence that both pathways respond to depicted objects by attributing salient emotional properties concerning both recognition and (apparent) action (Ferretti 2017a). The lateral orbitofrontal cortex and its related cortical projections related to the ventral pathway respond to pictorial contents with high emotional value (Wright et al. 2004; Schienle et al. 2006; Phillips et al. 2000; Shapira et al. 2003). Concerning the orbitofrontal/dorsal link, depictions that represent graspable objects that would be, in principle, approachable without any risk, activate a facilitating motor response, while depicted dangerous objects evoke a motor resonance concerning aversive action possibilities that recalls an interference-effect between the information about the object's potential risks and the action possibilities detected (Anelli et al. 2012; for a complete review see Ferretti 2017a: 5.1).

These results suggest that depicted objects do, indeed, induce an equal behavior of our visual system as face-to-face perception does. We can equally attribute several properties to both of these perceptual situations. Thus, this evidence supports the stance used by neuroscientists, who are justified in their everyday practice with pictures: a depicted object—which is just one of the two components of a picture—and a normal object are, in this sense, very similar objects of perception.

The equal attribution of these specific properties is possible even if, unlike normal objects, depicted objects cannot foster any visual feeling of presence (section 2). The representation of the difference concerning presence is mainly due to the computational behavior of the ventral stream, which detects, at least at the unconscious level of visual processing, the presence of the surface (see point (b)

https://doi.org/10.1017/apa.2017.28 Published online by Cambridge University Press

above). This is in line with the fact that we usually focus on the depiction, while unconsciously representing the surface.

But the fact that depicted objects induce an equal behavior of our visual system with respect to the properties mentioned above, even if they do not look present, is sufficient in order to rely on experimental settings that make use of depicted objects rather than normal objects. Indeed, if our visual system can attribute, say, a color property, a shape property, or the like to both a normal and a depicted object, the neuroscientific practice of using depicted objects when aiming at discovering the neural correlates of vision is reliable. This supports the *Sharpened Weak* NT.

4. Conclusion: The Pictorial World: Between Philosophers' and Neuroscientists' 'Pictures'

There are equally good reasons to support both PT and the *Sharpened Weak* NT. Indeed, the evidence in favor of the *Sharpened Weak* NT does not conflict with PT. In order to clearly appreciate this point, we can come back to the perceptual hierarchy mentioned in section 3.

Face-to-face perception and the perception of depicted objects allow a very similar property attribution. This makes a depicted object and a normal object very similar objects of perception, in line with the *Sharpened Weak* NT. This remains true even if only the former can foster the visual feeling of presence. Hence, face-to-face and picture perception are not exactly the same perceptual phenomenon. This is in virtue of the fact that, in line with PT, the perception of a depicted object is a 'perceptual part' of a more complex phenomenon: picture perception. And picture perception is different from face-to-face perception: it involves two representations—— the *Double Representation Thesis*. This difference is captured by PT and in line with the *Sharpened Weak* NT. Moreover, in picture perception it is the representation of the surface as present that prevents us from representing the depicted object as present. The *Weak* NT does not mention this perceptual fact. And the *Strong* NT does not specify the crucial difference between a *depicted object* and a *picture*. Only the *Sharpened Weak* NT can account for all these specifications. Therefore, it is perfectly in line with PT. But these differences between the perception of depicted objects and face-to-face perception do not rule out their similarity concerning the perceptual attribution of several properties (e.g., shape, geometry, slant, depth, color, action properties, etc.). Furthermore, the resemblances between face-to-face and picture perception, invoked by PT, are granted by the resemblances between face-to-face perception and the perception of depicted objects invoked by the *Sharpened Weak* NT. This explanation saves both the *Sharpened Weak* NT and PT and is perfectly in line with our most recent and satisfying accounts of picture perception (Nanay 2017).

Summing up, the regimentation offered here explains why the stance of philosophers and that of neuroscientists toward picture perception, even if apparently in conflict, can coexist without any problem. Therefore, neuroscientists are justified in using depicted objects as substitutes of normal objects in their

https://doi.org/10.1017/apa.2017.28 Published online by Cambridge University Press

experimental settings, and philosophers are justified in arguing for the peculiar status of pictures as particular objects of perception.

GABRIELE FERRETTI
UNIVERSITY OF FLORENCE
gabriele.ferretti88@gmail.com

References

Aasen, S. (2015) 'Pictures, Presence and Visibility'. *Philosophical Studies*, 187–203. doi:10.1007/s11098-015-0475-4.
Abell, C. (2009) 'Canny Resemblance'. *Philosophical Review*, 118, 183–223.
Abell, C., and K. Bantinaki, eds. (2010) *Philosophical Perspectives on Depiction*. Oxford: Oxford University Press.
Anelli, F., A. M. Borghi, and R. Nicoletti. (2012) 'Grasping the Pain: Motor Resonance with Dangerous Affordances'. *Consciousness and Cognition*, 21, 1627–39.
Barrett, L. F., and L. F. Bar. (2009) 'See it with Feeling: Affective Predictions during Object Perception'. *Philosophical Transactions of the Royal Society*, 364, 1325–34.
Barry, S. (2009) *Fixing My Gaze*. New York: Basic Books.
Berryhill, M. E., and I. R. Olson. (2009) 'The Representation of Object Distance: Evidence from Neuroimaging and Neuropsychology'. *Frontiers in Human Neuroscience*, 3, 1–9.
Briscoe, R. (2009) 'Egocentric Spatial Representation in Action and Perception'. *Philosophy and Phenomenological Research*, 79, 423–60.
Briscoe, R. (2016) 'Depiction, Pictorial Experience, and Vision Science'. In C. Hill and B. McLaughlin (eds.), *Philosophical Topics, Special Issue on Appearance Properties*, 44, 43–81.
Briscoe, R., and J. Schwenkler. (2015) 'Conscious Vision in Action'. *Cognitive Science*, 39, 1435–67.
Bruzzo, A., A. M. Borghi, and S. Ghirlanda. (2008) 'Hand-object Interaction in Perspective'. *Neuroscience Letters*, 441, 61–65.
Buccino, G., M. Sato, L. Cattaneo, F. Rodà, and L. Riggio. (2009) 'Broken Affordances, Broken Objects: A TMS Study'. *Neuropsychologia*, 47, 3074–78.
Castiello, U. (2005) 'The Neuroscience of Grasping'. *Nature Reviews*, 6, 726–36.
Cavedon-Taylor, D. (2011) 'The Space of Seeing-In'. *British Journal of Aesthetics*, 51, 271–78.
Chao, L. L., and A. Martin. (2000) 'Representation of Manipulable Man-made Objects in the Dorsal Stream'. *NeuroImage*, 12, 478–84.
Chasid, A. (2014) 'Pictorial Experience: Not so Special after all'. *Philosophical Studies*, 171, 471–91.
Chinellato, E., and A. P. del Pobil. (2016) *The Visual Neuroscience of Robotic Grasping: Achieving Sensorimotor Skills through Dorsal-ventral Stream Integration*. Zurich, Switzerland: Springer International Publishing.
Costantini, M., E. Ambrosini, G. Tieri, C. Sinigaglia, and G. Committeri. (2010) 'Where Does an Object Trigger an Action? An Investigation about Affordance in Space'. *Experimental Brain Research*, 207, 95–103.
Cutting, J. (2003) 'Reconceiving Perceptual Space'. In H. Hecht, R. Schwartz, and M. Atherton (eds.), *Looking into Pictures: An Interdisciplinary Approach to Pictorial Space* (Cambridge, MA: MIT Press), 215–38.
Dijkerman, H. C., R. D. McIntosh, I. Schindler, T. C. W. Nijboer, and A. D. Milner. (2009) 'Choosing between Alternative Wrist Postures: Action Planning Needs Perception'. *Neuropsychologia*, 47, 1476–82.

https://doi.org/10.1017/apa.2017.28 Published online by Cambridge University Press

Elliott, R., R. J. Dolan, and C. D. Frith. (2000) 'Dissociable Functions in the Medial and Lateral Orbitofrontal Cortex: Evidence from Human Neuroimaging Studies'. *Cerebral Cortex*, 10, 308–17. doi:10.1093/cercor/10.3.308.

Ferretti, G. (2016a) 'Pictures, Action Properties and Motor-related Effects'. *Synthese Special Issue Neuroscience and Its Philosophy*, 193, 3787–817. doi:10.1007/s11229-016-1097-x.

Ferretti, G. (2016b) 'Through the Forest of Motor Representations'. *Consciousness and Cognition*, 43, 177–96. doi:10.1016/j.concog.2016.05.013.

Ferretti, G. (2016c) 'Visual Feeling of Presence.' *Pacific Philosophical Quarterly*. doi:10.1111/papq.12170.

Ferretti, G. (2017a) 'Pictures, Emotions, and the Dorsal/Ventral account of Picture Perception'. *Review of Philosophy and Psychology*, 8, 595–616. doi:10.1007/s13164-017-0330-y.

Ferretti, G. (2017b). 'Two Visual Systems in Molyneux Subjects'. *Phenomenology and the Cognitive Sciences*, 1–37. https://doi.org/10.1007/s11097-017-9533-z.

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

Gonzalez, C., T. Ganel, R. Whitwell, B. Morrissey, and M. Goodale. (2008) 'Practice Makes Perfect, but Only with the Right Hand: Sensitivity to Perceptual Illusions with Awkward Grasps Decreases with Practice in the Right, but not the Left Hand'. *Neuropsychologia*, 46, 624–31.

Goodale, M. A., and A. D. Milner. (2004) *Sight Unseen*. Oxford: Oxford University Press.

Gregory, R. L., (2012) 'Pictures as Strange Objects of Perception'. In F. G. Barth, P. Giampieri-Deutsch, and H.-D. Klein (eds.), *Sensory Perception: Mind and Matter* (Amsterdam: Springer), 175–81.

Grézes, J., and J. Decety. (2002) 'Does Visual Perception of Object Afford Action? Evidence from a Neuroimaging Study'. *Neuropsychologia*, 40, 212–22.

Grill-Spector, K. and K. S. Weiner. (2014) 'The Functional Architecture of the Ventral Temporal Cortex and its Role in Categorization'. *Nature Reviews, Neuroscience*, 15, 536–48 doi:10.1038/nrn3747.

Grush, R. (2000) 'Self, World, and Space: The Meaning and Mechanisms of Ego- and Allocentric Spatial Representation'. *Brain and Mind*, 1, 59–92.

Hecht, H., R. Schwartz, and M. Atherton, eds. (2003) *Looking into Pictures: An Interdisciplinary Approach to Pictorial Space*. Cambridge, MA: MIT Press.

Hopkins, R. (2010) 'Inflected Pictorial Experience: Its Treatment and Significance'. In C. Abell and K. Bantinaki (eds.), *Philosophical Perspectives on Depiction* (Oxford: Oxford University Press), 151–80.

Hopkins, R. (2012) 'Seeing-in and Seeming to See'. *Analysis*, 72, 650–59.

Jacob, P., and M. Jeannerod. (2003) *Ways of Seeing: The Scope and Limits of Visual Cognition*. Oxford: Oxford University Press.

James, T., G. Humphrey, J. Gati, R. Menon, and M. Goodale. (2002) 'Differential Effects of Viewpoint on Object-driven Activation in Dorsal and Ventral Stream'. *Neuron*, 35, 793–801.

Jeannerod, M., J. Decety, and F. Michel. (1994) 'Impairment of Grasping Movements Following a Bilateral Posterior Parietal Lesion'. *Neuropsychologia*, 32, 369–380.

Koenderink, J. J. (2012) *Pictorial Space*. Utrecht: De Clootcrans Press.

Koenderink, J. J., and A. J. van Doorn. (2003) 'Pictorial Space'. In H. Hecht, R. Schwartz, and M. Atherton (eds.), *Looking into Pictures: An Interdisciplinary Approach to Pictorial Space* (Cambridge, MA: MIT Press), 215–38.

Kulvicki, J. (2006) *On Images*. Oxford: Clarendon Press.

Levinson, J. (1998) 'Wollheim On Pictorial Representation'. *Journal of Aesthetics and Art Criticism*, 56, 227–33.

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

Lopes, D.M. (2005) *Sight and Sensibility: Evaluating Pictures*. Oxford: Oxford University Press.

Marotta, J., and M. Goodale. (2001) 'The Role of Familiar Size in the Control of Grasping'. *Journal of Cognitive Neuroscience*, 13, 8–17.

Matthen, M. (2005) *Seeing, Doing and Knowing: A Philosophical Theory of Sense Perception*. Oxford: Oxford University Press.

https://doi.org/10.1017/apa.2017.28 Published online by Cambridge University Press

Milner, A. D., and M. A. Goodale. (1995]2006) *The Visual Brain in Action*. 2nd ed. Oxford: Oxford University Press.
Milner, A. D., and M. A. Goodale. (2008) 'Two Visual Systems Re-viewed'. *Neuropsychologia*, 46, 774–85.
Murata, A., V. Gallese, G. Luppino, M. Kaseda, and H. Sakata. (2000) 'Selectivity for the Shape, Size and Orientation of Objects for Grasping in Neurons of Monkey Parietal Area AIP.' *Journal of Neurophisiology*, 79, 2580–2601.
Nanay, B. (2010a) 'Transparency and Sensorimotor Contingencies: Do We See through Photographs?' *Pacific Philosophical Quarterly*, 91, 463–80.
Nanay, B. (2010b) 'Inflected and Uninflected Experience of Pictures'. In C. Abell and K. Bantilaki (eds.), *Philosophical Perspectives on Depiction* (Oxford: Oxford University Press), 181–207.
Nanay, B. (2011) 'Perceiving Pictures'. *Phenomenology and the Cognitive Sciences*, 10, 461–80.
Nanay, B. (2013) *Between Perception and Action*. Oxford: Oxford University Press.
Nanay, B. (2015) 'Trompe l'oeil and the Dorsal/ventral Account of Picture Perception'. *Review of Philosophy and Psychology*, 6, 181–97.
Nanay, B. (2016) *Aesthetics as Philosophy of Perception*. Oxford: Oxford University Press.
Nanay, B. (2017) 'Threefoldness'. *Philosophical Studies*, 1–20. doi:10.1007/s11098-017-0860-2.
Nelissen, K., O. Joly, J. B. Durand, J. T. Todd, W. Vanduffel, and G. A. Orban. (2009) 'The Extraction of Depth Structure from Shading and Texture in the Macaque Brain.' *PloS One*, 4(12), e8306. doi:10.1371/journal.pone.0008306.
Noë, A. (2004) *Action in Perception*. Cambridge, MA: MIT Press.
O'Reilly, R. C. (2010) 'The What and How of Prefrontal Cortical Organization'. *Trends in Neurosciences*, 33, 355–61.
Pani, P., T. Theys, M. C. Romero, and P. Janssen. (2014) 'Grasping Execution and Grasping Observation Activity of Single Neurons in the Macaque Anterior Intraparietal Area'. *Journal of Cognitive Neuroscience*, 26, 2342–55.
Phillips, M. L., I. M. Marks, C. Senior, D. Lythgoe, A. M. O'Dwyer, O. Meehan, et al. (2000) 'A Differential Neural Response in Obsessive-compulsive Disorder Patients with Washing Compared with Checking Symptoms to Disgust.' *Psychological Medicine*, 30, 1037–50.
Proverbio, M. A., R. Adorni, and G. E. D'Aniello. (2011) '250 ms to Code for Action Affordance During Observation of Manipulable Objects'. *Neuropsychologia*, 49, 2711–19.
Raos, V., M. A. Umilta, A. Murata, L. Fogassi, and, V. Gallese. (2006) 'Functional Properties of Grasping-related Neurons in the Ventral Premotor Area F5 of the Macaque Monkey'. *Journal of Neurophysiology*, 95, 709–29.
Romero, M. C., P. Pani, and P. Janssen. (2014) 'Coding of Shape features in the macaque anterior intraparietal area systems/circuits.' *The Journal of Neuroscience*, 34(11), 4006–4021.
Sakata, H., M. Taira, A. Murata, and S. Mine. (1995) 'Neural Mechanisms of Visual Guidance of Hand Action in the Parietal Cortex of the Monkey'. *Cerebral Cortex*, 5, 429–38.
Sakata, H., K. Tsutsui, and M. Taira. (2003) Representation of the 3D world in art and in the brain., *International Congress Series*, 1250, 5–35.
Schienle, A., A. Schäfer, A. Hermann, B. Walter, R. Stark, and D. Vaitl. (2006) 'fMRI Responses to Pictures of Mutilation and Contamination'. *Neuroscience Letters*, 393, 174–78.
Schwartz, R., and M. Atherton, eds. (2003). *Looking into Pictures: An Interdisciplinary Approach to Pictorial Space*. Cambridge, MA: MIT Press.
Schwenkler, J. (2013) 'Do Things Look the Way they Feel?' *Analysis*, 73, 86–96.
Shapira, N.A., Y. Liu, A.G. He, M. M. Bradley, M. C. Lessig, G. A. James, D. J. Stein, P. J. Lang, and W. K. Goodman. (2003) 'Brain Activation by Disgust-inducing Pictures in Obsessive-compulsive Disorder'. *Biological Psychiatry*, 54, 751–56.
Siegel, S. (2007) 'How Can We Discover the Contents of Experience?' *Southern Journal of Philosophy* (Supp.), 45, 127–42.
Siegel, S. (2010) *The Contents of Visual Experience*. New York: Oxford University Press.
Snow, J. C., C. E. Pettypiece, T. D. MacAdam, A. D. McLean, P. W. Stroman, M. A. Goodale, and J. C. Culham. (2011) 'Bringing the Real World into the fMRI Scanner: Repetition Effects for Pictures versus Real Objects'. *Scientific Reports*, 1. doi: 10.1038/srep00130.

https://doi.org/10.1017/apa.2017.28 Published online by Cambridge University Press

Taira, M., I. Nose, K. Inoue, and K. Tsutsui. (2001) 'Cortical Areas Related to Attention to 3D Surface.' *Neuroimage*, 14, 959–66.
Turnbull, O. H., Driver, J., and R. A. McCarthy. (2004) '2D but not 3D: Pictorial Depth Deficits in a Case of Visual Agnosia.' *Cortex*, 40, 723–38.
Vishwanath, D. (2011). 'Information in Surface and Depth Perception: Reconciling Pictures and Reality'. In L. Albertazzi, G. J. van Tonder, and D. Vishwanath (eds.), *Perception Beyond Inference: The Information Content of Visual Processes* (Cambridge, MA: MIT Press), 201–40.
Vishwanath, D. (2014) 'Toward a New Theory of Stereopsis'. *Psychological Review*, 121, 151–78.
Vishwanath, D., and P. Hibbard. (2010) 'Quality in Depth Perception: The Plastic Effect'. *Journal of Vision*, 10, 1673–85.
Vishwanath, D., and P. Hibbard. (2013) 'Seeing in 3D With Just One Eye: Stereopsis in the Absence of Binocular Disparities'. *Psychological Science*, 24, 1673–85.
Vishwanath, D., A. R. Girshick, andM. S. Banks. (2005) 'Why Pictures look Right when Viewed from the Wrong Place'. *Nature Neuroscience*, 8, 1401–10.
Voltolini, A. (2013) 'Why, As Responsible for Figurativity, Seeing-In Can Only be Inflected Seeing-In'. *Phenomenology and the Cognitive Sciences*, 14, 651–67. doi:10.1007/s11097-013-9335-x.
Westwood, D., J. Danckert, P. Servos, and M. Goodale. (2002) 'Grasping two-dimensional Images and three-dimensional Objects in visual-form Agnosia'. *Experimental Brain Research*, 144, 262–67.
Wollheim, R. (1980) 'Seeing-As, Seeing-In, and Pictorial Representation'. In R. Wollheim (ed.), *Art and its Object*, 2nd ed. (Cambridge: Cambridge University Press), 205–26.
Wollheim, R. (1998) 'On Pictorial Representation'. *Journal of Aesthetics and Art Criticism*, 56: 217–26.
Wright, P., G. He, N. A. Shapira, W. K. Goodman, and Y. Liu. (2004) 'Disgust and the Insula: fMRI Responses to Pictures of Mutilation and Contamination'. *Neuroreport*, 15, 2347–51.
Zipoli Caiani, S. (2013) 'Extending the notion of affordance.' *Phenomenology and the Cognitive Sciences*, 13, 275–93.
Zipoli Caiani, S., and G. Ferretti. (2017) 'Semantic and Pragmatic Integration in Vision for Action'. *Consciousness and Cognition*, 48, 40–54.

https://doi.org/10.1017/apa.2017.28 Published online by Cambridge University Press