Depiction, Pictorial Experience, and Vision Science

Robert Briscoe
Ohio University

ABSTRACT. Pictures are patterned, 2D surfaces designed to elicit 3D-scene-representing experiences from their viewers. In this essay, I argue that philosophers have tended to underestimate the relevance of research in vision science to understanding the nature of pictorial experience. Both the deeply entrenched methodology of virtual psychophysics as well as empirical studies of pictorial space perception provide compelling support for the view that pictorial experience and seeing face-to-face are experiences of the same psychological, explanatory kind. I also show that an empirically informed account of pictorial experience provides resources to develop a novel, resemblance-based account of depiction. According to what I call the *deep resemblance theory*, pictures work by presenting virtual models of objects and scenes in phenomenally 3D, pictorial space.

> *Most people think they know what a picture is. Anything so familiar*
> *must be simple. They are wrong.*
> —*Gibson 1980, xvii*

A picture is a patterned, 2D surface that, when present to sight, displays the appearance of an absent, three-dimensionally organized scene. Consider the experience elicited by Pieter Bruegel the Elder's painting *The Hunters in the Snow* (fig. 1). The objects that we experience *in* the painting—on the far side of the pictorial surface,

as it were—appear to have different voluminous, 3D shapes, but the painting itself appears unambiguously flat. The objects that we experience in the painting appear to recede from the surface, but the painting itself appears to be positioned on a single plane of depth. The objects that we experience in the painting appear to be illuminated from above, but the painting itself may appear to be illuminated from the side or some other direction. The main point, for present purposes, is that the range of appearance properties—the shapes, sizes, colors, textures, orientations, and so forth—attributed by pictorial experience to its objects typically contrasts with those that are actually exemplified by the surface before our eyes.

Following standard usage in art history, aesthetics, and cognitive science, I shall refer to the 3D-scene-representing experience elicited by a 2D, pictorial surface as the experience of "pictorial space" (Wölfflin 1929; Pirenne 1970; Kubovy 1986; Rogers 1995, 2003; Koenderink 1998, 2012; Hecht et al. 2003; Thompson et al. 2011, chap. 12). It is a fundamental challenge for philosophical theories of depiction to provide an adequate account of the nature of this experience and to specify its role, if any, in supporting the distinctive function that pictures perform. The resemblance theory of depiction developed in this essay pursues a novel approach to meeting this challenge. In particular, it disputes two seemingly self-evident assumptions that have structured recent work on depiction.

First, philosophical discussions of picturing and pictorial experience are typically structured by the assumption that depiction is a genus of *representation* (Goodman 1976; Wollheim 1987, 1998, 2003; Schier 1986; Walton 1990, 2008; Haugeland 1991; Budd 1992/2003, 1993/2003; Lopes 1996, 2005, 2006; Hopkins 1998, 2006; Abell 2009; Newall 2009; Nanay 2010; Greenberg 2013). Pictures, Malcolm Budd, for instance, writes:

> are a distinct kind of representation: it is definitive of a picture that it represents what it depicts by depicting it, and depiction is a form of representation different from any other. (Budd 1993/2003, 216)

And here is Robert Hopkins:

> [W]hen we encounter pictures we encounter objects that are patently communicative episodes. They are messages in a bottle of the visual world. (Hopkins 1998, 156)

And Dominic Lopes:

> [P]ictures are at bottom vehicles for the storage, manipulation, and communication of information . . . Pictures share language's burden in representing the world and our thoughts about it. (Lopes 1996, 7)

> Depiction . . . is a form of representation. This is one of the few bedrock truths approved by all philosophers who have worked up an opinion on the matter. (Lopes 2006, 160)

We can call the general, but seldom, if ever, defended, assumption common to these statements the (Pictorial) Content Thesis. According to the Content Thesis, pictures, like assertions, beliefs, and perceptual experiences, are constitutively representational:

> If a surface *P* depicts an object *O* then, necessarily, there is some repre-
> sentational content *C* such that *P* has the function of communicating
> *C* to its intended viewers. This is part of *what it is* for *P* to function as
> a picture of *O*.[1]

Hence, a central task of an account of depiction is to explain which properties of pictorial representational systems distinguish them from nonpictorial representational systems, in particular, representation by means of language (Haugeland 1991).

Proponents of the Content Thesis also normally approve as a "bedrock truth" what we can call the (Pictorial) Vehicle Thesis:

> The vehicle of a picture's representational content is the pattern or
> design visible on the picture's 2D surface.

The vehicles of pictorial representational content, like the vehicles of written, linguistic representational content, on this view, are superficial marks or patterns, for example, an arrangement of paint strokes on a canvas or an array of light emitting pixels in a laptop display. In consequence, advocates of the Vehicle Thesis typically assume as self-evident that our ability to work out a picture's representational content is explanatorily dependent on seeing the picture's design properties, that is, on visually perceiving the structure or organization of a 2D pattern on the pictorial surface and understanding its representational significance. Differences in the apparent structure or organization of the picture's design are supposed to explain differences in how suitably equipped viewers interpret or make sense of the picture. Understanding a painting, drawing, or photograph of a scene is, in this respect, if not others, like reading a written description of it.

The empirically informed resemblance theory of depiction defended in this essay rejects both structuring assumptions. Here is an overview of the argument to follow. In section 1, I begin by discussing real-world models and their use. A *model*, as stipulatively understood here, is an artifact designed to simulate the outward visual appearance of an actual or possible object or scene—the model's *original*. While some models perform a genuinely representational function, others, I argue, perform a function that is rather substitutive in nature. They do not stand for their originals, but rather stand in for them, in a sense to be explained. Building on work by Ruth Millikan, I then suggest two necessary conditions for iconic representation by means of modeling. In section 2, I argue that philosophers have tended to underestimate the relevance of research in vision science to understanding the nature of pictorial experience. I then show that an empirically informed account of pictorial experience provides resources to develop a novel, resemblance-based account of depiction. According to what I call the deep resemblance theory, pictures work

---

1. The Content Thesis is formulated here in a general way that doesn't presuppose any particular account of pictorial content. For some widely accepted constraints on an adequate account of pictorial content, see Hopkins 2006, 145–46. For an expression of the view that "pictures are human artefacts specifically designed for communication," see Lopes 1996, 19, 86–89.

by displaying virtual models of objects and scenes in phenomenally 3D, pictorial space. Virtual models, to borrow M. G. F. Martin's (2012) phrase, are *mere visibilia*: they "[present] the appearance of a material object to the viewer without . . . exemplifying that appearance" (2012, 344). I argue that the distinction between representational and substitutive models also applies to models constructed in pictorial space and offer two necessary conditions for iconic representation by means of pictorial modeling (these are analogues of the conditions for real-world models). Drawing on resources provided by the deep resemblance theory, I subsequently explain why I think that both the Content and Vehicle Thesis are false. In section 3, I conclude by briefly discussing some of the more interesting consequences of the deep resemblance theory.

1. REAL-WORLD MODELS

Central to the deep resemblance theory is the idea that pictures function by presenting *virtual models* of objects and scenes in pictorial space. In order to motivate this proposal and to draw out its consequences, it is helpful to begin by reflecting on real-world models and their use. I stipulatively use the term "model" here in a broad way that includes 3D scale-models and sculptures, but also extends to decoys, dummies, dioramas, mock-ups, and movie set façades. Technical definitions of "model" in mathematics, logic, and the philosophy of science are not intended. They should be kept at arm's length in what follows.[2]

Real-world models, as understood here, are distinguished from other mimetic artifacts by two main features. First, models are designed to simulate outward visual appearances or "looks." A model may or may not be designed to share certain functional or material properties with its original, but it must be designed to share certain visual appearance properties with it. Some models, for example, natural history museum dioramas, are exquisitely detailed. Others are highly abstract or stylized. The relative positions of wooden blocks on a table, for example, can be used to model the 3D layout of objects in a real-world scene. Which correspondences are set up between a model and its original depends on the model's intended function, on how its users are supposed to interpret or interact with it. (Compare, in this connection, the role of provisional study models in the architectural design process with the role of intricate engineering or construction models.)

---

2. In the philosophy of science, in particular, models are usually thought of as concrete, computational, or mathematical structures whose properties stand in representational relations to properties of a target system. For an illuminating discussion, see Weisberg 2013. My use of the term "model" here doesn't presuppose that models are vehicles of representation. That a given model, on the account to be developed below, functions as a representation of an object or scene is a contingent, empirical fact about how the model is used and not something constitutive of modeling in general.

For present purposes, however, the main point is that a model is always contrived to capture the visual appearance of its original or aspects thereof. In consequence, the properties that a model must purposefully share with its original are visible properties. These include both properties registered by low-level feature detectors such as shape, size, color, texture, and albedo as well as "gestalt" properties built-up from the former, e.g., a certain color scheme or arrangement of shapes.

Second, models do not typically belong to the same artifactual or natural kind as their original. A scale-model of a cathedral isn't a cathedral; the funerary effigy of Elizabeth I in Westminster Abbey isn't a deceased queen; a stone lion isn't a lion; a decoy duck isn't a duck; and so on. Hans Kamp (1975) defines a privative adjective as one for which, given an adjective *A* and noun *N*, the assertion "*AN* is a *N*" is never true. E.g., the assertion "the stone lion is a lion" or "the dummy tank is a tank" is always false. A privative concept, by extension, is a concept that combines with other concepts at the level of thought in a manner analogous to a privative adjective (Franks 1995). Although much ink has been spilled over the generality of Kamp's analysis—"a 'stone lion' may be a very stoic lion, or a lion which stands very still" (Coulson and Fauconnier 1999)—it usefully applies to the concept of a model as understood here. The result of combining the concept *model* or one of its associated subconcepts, e.g., *sculpture* or *dummy*, with the concept of a natural or artifactual kind *F*, e.g., *lion* or *tank*, is a combined concept that typically excludes actual *Fs* from its extension. No stone lion is a lion. No dummy tank is a tank.

Many models have an overtly representational, property-attributing function. For present purposes, I assume a fairly standard conception of representation in philosophy and cognitive science:

> [A] representation is one thing that is taken to stand for another, in a way relevant to the control of behavior or some other decision. More specifically, I take the paradigm case here to be that when a person decides to control their behavior towards one domain, *Y*, by attending to the state of something else, *X*. The state of *X* is "consulted" in working out how to behave in relation to *Y*. (Godfrey-Smith 2006, 45)

Models that represent their originals by means of an intention-based overlap in visual appearance properties are iconic representations (or *icons* for short). The distinguishing mark of an icon is the presence of what Ruth Millikan (2004) refers to as "reflexivity." A representation is reflexive when the instantiation of a determinate property by the representation's vehicle signifies instantiation of that property in the represented domain. The colors and shapes of the doors in an architectural model, for example, can be used to represent the colors and shapes of the doors in a real-world house. An icon embodies what Millikan calls a "relative reflexive" element when the instantiation of a determinate property by its vehicle signifies the instantiation of another determinate in the same determinable range by means of a preestablished mapping function. Thus, "if one inch on a blueprint stands for one inch, length is a reflexive element of the blueprint sign. If one inch stands for one foot, length is a relative reflexive" (Millikan 2004, 53).

Icons are excellent devices for storing and communicating information about an object's visual appearance. Indeed, an icon that veridically represents the visible properties of its original enables its viewers vicariously to enjoy an experience as of those properties. When accurate and inspected under appropriate viewing conditions, an icon elicits an experience that is, we might say, *transparent* to the visible properties of its original. In contrast with a veridical, verbal description of an object *O*, a veridical iconic representation of *O* enables its viewers to form perceptually based beliefs about *O*'s properties (and, in some cases, to rehearse visuomotor actions sensitive to those properties) without seeing *O* itself.

A model *M* of an object *O*, however, need not function as an iconic representation of *O*. Rather, in many cases, intended resemblances between *M* and *O* are required not in order that *M* reflexively represent *O* as having certain properties, but rather in order that *M* perform some role or achieve some effect performed or achieved by *O*. If the visual appearance of a counterfeit $100 bill, for instance, doesn't sufficiently resemble that of a real $100 bill, then the counterfeit will not perform the role for which real $100 bills are typically tendered. Similarly, if the visual appearance of a decoy duck doesn't sufficiently resemble that of a real duck, its presence on a pond will not have the effect of luring other ducks to the vicinity. Neither the counterfeit nor the decoy, when used for the purposes for which it is standardly produced, however, is an iconic representation. Neither has a communicative, property-attributing function. In general, intended visual resemblance is necessary, but *not sufficient* for representation by means of modeling. Whether or not a model is an icon depends on the way that it is used, on the kind of work that is done with it.

Certain models, I have suggested, have a function that is broadly substitutive rather than representational in nature. They do not stand for their originals in the sense of representing or attributing properties to them. Instead, they are designed to *stand in* for them. Substitutive models are by no means rare and perform a diverse range of functions. Some substitutive models, as already noted, are designed to deceive or delude (dummies, lures, and decoys). Others perform an apotropaic function (the stone lions that "guard" many Buddhist temples in East Asia). Some are intended to achieve certain ornamental or aesthetic effects (the acanthus leaves that adorn Corinthian columns or silk flowers on a restaurant table). Others are designed to be objects of pretend-play (dolls and doll houses).

In the present context, the main point is that models, in the sense introduced here, do not always have a representational, property-attributing function. Models aren't always icons. Minimally, I would propose that a model *M* of an original *O* must meet two Millikanian requirements in order to represent *O* iconically:

> *Guidance*
> M must have the function of guiding the way its viewers perform some task *T*, where *T* involves either engaging in actions directed in relation to *O* or forming perceptually based beliefs about *O*'s visible attributes, and, further, it is a condition of the (nonaccidental) successful performance of *T* that *M* actually resemble *O* in intended, visible respects.

Systematicity
The way viewers are guided by *M* in performing task *T* must systematically depend on *M*'s visible attributes, such that had the intention-based respects in which *M* visually resembles *O* been different, then viewers' performance of *T* would have proved successful only if *O*'s visible attributes had been different in corresponding respects.

The Guidance Requirement is an expression of the familiar idea that representation is a *functional* kind, in particular, that what makes something a representation is the role it plays in adapting the activities of its receivers or, as Millikan would say, its "consumers" to the environment (Millikan 1989, 2004). A model has the functional status or role of an iconic representation, according to the requirement, only if *used* by its viewers in certain ways. A model that has neither the job of guiding the way its viewers form beliefs about the visible attributes of its original *O*, nor the job of guiding the way its viewers perform actions directed in relation to *O* may have various other useful purposes, but it doesn't have the job of iconically representing *O*.

The point of the Systematicity Requirement is that representations always belong to systems of representation. Transformations that alter how a given representation is constructed or articulated must be possible (a representation, as Millikan puts it, must have significant variable aspects or parts), and these must give rise, in a regular way, to corresponding differences in how its users are normally guided by the representation in their cognitive and practical undertakings.

Substitutive models do not meet these requirements. The visible attributes of a counterfeit $100 bill nonaccidentally match, but do not reflexively represent the visible attributes of a real $100 bill because the counterfeit's function is not to guide the way its viewers either interact with or form beliefs about the visible attributes of real $100 bills. For similar reasons, the visible attributes of a decoy duck do not reflexively represent those of a real duck.

That a given model iconically represents its original, then, is a contingent, empirical fact about how the model is used and not something that is constitutive of modeling in general. In order to prevent misunderstanding of this claim, two further points about real-world models need to be made at this junction.

1.1. STANDARDS OF CORRECTNESS FOR MODELS

First, where there is representation, there is the possibility of error, that is, misrepresentation. Assertions, beliefs, and experiences have contents that are assessable for veridicality: a representation *R* is veridical when the world is the way *R* represents it as being and nonveridical otherwise. From the standpoint of psychological explanation, there is representation only where there is need to make nontrivial appeal to states (items, patterns, events, etc.) with veridicality conditions (for discussion, see Burge 2010).

Substitutive models—and this is the first point—have correctness conditions that can easily be mistaken for representational veridicality conditions. Models, by

definition, are intended to resemble their originals in certain visually accessible respects. Suppose that an object O is red, round, and shiny and that a substitutive model M is intended to resemble O in respect of color, shape, and sheen. This then sets up a standard of correctness to which M may fail to conform. M, for example, may fail sufficiently to match O in color. If so, then the model-maker's mimetic intention will not be fulfilled. This, by itself, however, doesn't mean that M misrepresents O's color, i.e., nonveridically attributes a certain color to O. Failure to fulfill the relevant mimetic intention doesn't entail misrepresentation if M's function is substitutive rather than the representational—if M's job is to serve as something that stands in for (rather than stands for) its original.[3]

This points to an additional standard of correctness for substitutive models: a substitutive model M must correspond to its original O in certain respects if M is to perform successfully its intended *function*, whatever that happens to be. That a model M satisfies the first standard of correctness doesn't by itself entail that M will correspond to O in those respects necessary for M to serve successfully as a proxy or stand-in for O. The particular respects, e.g., in which a lure has been designed visually to resemble a small fish may not correspond to those that would, in fact, reliably cause bigger fish to attempt to swallow it. The possibility of such failure in correspondence, however, again doesn't amount to the possibility of misrepresentation if the model's function is substitutive rather than representational.

1.2. MODELS AND NATURAL INFORMATION

Models typically carry *natural information* about the visible properties of their originals. It is important, however—and this is the second point—not to confuse the presence of natural information with the presence of intentional representation. Most basically, a thing a conveys natural information about another thing b, whenever there is some real connection between them in nature such that you can use a to learn about b (Millikan 2004, 2012). The causal connection between the depth of a bear's paw print and the bear's weight is such a real connection in nature. So is the correlation between the number of rings in a tree trunk and the tree's age. And so is the purposeful correspondence between a model M and its original O: if M's visible attributes nonaccidentally overlap with O's, then one can learn about O's visual appearance—and, perhaps, even learn to recognize O—by viewing M. That M's appearance carries natural information about O's appearance, however, doesn't entail that M's visual attributes reflexively represent O's visual attributes. You can learn to recognize a duck by viewing a well-executed decoy, e.g., but the decoy's function *qua* decoy is substitutive rather than representational.

---

3. A point also made by M. G. F. Martin: "When we ask of the forgery of a Ming vase whether it is an accurate copy of the original, we do not assume that either the original or the copy possesses a representational content. So in general, to ask of something whether it is accurate or not need not require it to be a representation or to have representational content, even if in some specific cases it is; it is simply to invite someone to match things" (2010, 223).

It will be important to keep both of the preceding points in mind when evaluating the claim, presented in the next section, that virtual models constructed in pictorial space do not have an inherently representational function.

2. THE DEEP RESEMBLANCE THEORY

Pictures, according to resemblance theories of depiction, are akin to models: their job is to simulate the outward appearance of an actual or possible object of visual experience (Neander 1987; Budd 1992/2008, 1993/2008; Hopkins 1998, 2006; Abell 2009). For want of a better expression, I shall refer to such a pictorial simulation or likeness, in what follows, as a *pictorial model*. And I shall refer to the visual appearance properties in virtue of which a pictorial model is intended to resemble an object or scene as its resemblance properties (or *R-properties* for short).

Real-world models are 3D artifacts. The visual appearance properties in virtue of which a real-world model may resemble its original thus include voluminous shape and size. Pictures, by contrast, are typically flat or 2D. In consequence, it is natural to assume that the model a picture presents of its subject is experienced as *on* the pictorial surface and that R-properties are actual or experienced *design properties*. Most resemblance theories of depiction start from this assumption: they undertake to explain picturing in terms of an actual or experienced resemblance between a superficial likeness and an actual or possible object of visual experience. Hence, they face the challenge of specifying salient respects in which 2D pictorial surfaces resemble (or are experienced as resembling) the 3D objects and scenes that they depict (see Hopkins 1998, chap. 1).

2.1. PICTORIAL EXPERIENCE IN THE LIGHT OF VISION SCIENCE

There is another possibility, however. Instead of looking for R-properties on the pictorial surface, we can look for them *in* the pictorial surface, in what vision scientists and art historians refer to as "pictorial space." Consider again the experience elicited by *The Hunters in the Snow* (fig. 1). When we look at the painting, we do not simply experience a superficial pattern of colors on the canvas in front of us. We also experience a coherently organized, 3D scene. Some objects in the experienced scene appear relatively close to the pictorial point of view, while others, barely visible, appear in the far distance toward the line of the horizon. The space in which we represent these objects and their properties isn't the physical space through which we move our bodies and in which we locate the painting itself. Rather, to use J. J. Gibson's terminology, it is a "virtual space" constructed by the visual system in response to the collection of depth cues in the light reflected from the pictorial surface to the eye (1979, 281–83).

Importantly, whereas pictorial surfaces are typically 2D and perceived as such, objects in pictorial space are often experienced by human perceivers as robustly

Figure 1. Pieter Bruegel the Elder, *The Hunters in the Snow* (1565).

3D, exhibiting properties such as voluminous shape, surface curvature, orientation, interposition transparency, self-occlusion, and relative distance in depth (Ames 1925; Schlosberg 1951; Kennedy 1974; Rogers 1995, 2003; Cutting 2003; Koenderink & van Doorn 2003; Vishnawath 2011, 2014; Koenderink 1998, 2012). In consequence, the representational contents of pictorial experience significantly overlap with those of seeing face-to-face.

It is customary in discussions of depiction and pictorial experience to say that we "see" objects in pictures—for example, that we see dogs, trees, and some distant mountains in *The Hunters in the Snow*. Use of the verb "see" in this context, however, is potentially misleading. The reason is that the experience of seeing is factive. You can't see a dog unless there really *is* a dog visibly present in front of you. "Seeing," as Ryle would say, is a success term. By contrast, the experience of hallucinating a dog does not have this entailment. When you hallucinate, there is conscious visual representation *as of* an object with certain properties in the absence of an appropriate causal link to an object in the environment that actually exemplifies the properties you experience. In general, not all visual experience involves a relation to a *representatum* (for helpful discussion, see Burge 2010, 42–46). Just as there is nothing that you know when you believe falsely, there is nothing that you see when you hallucinate. Some visual representation is mere "visaging" to use a term coined by Ruth Millikan (2000).

Pictorial experience, on the account defended here, isn't factive. The objects of pictorial experience, strictly speaking, cannot be seen. They are, as Gibson puts it, merely "virtual objects." When you look, for example, at a picture of a dog, your experience represents an object as having a certain 3D shape, orientation, and position relative to other objects in the pictorial scene, even though there is no object in front of you that actually exemplifies any of these properties. In respect of its representational content, pictorial experience is thus doubly nonveridical: to use the terminology of Burge (2010), it involves failure both at the level of context-dependent, "singular applications" that function to single out particulars in the environment and at the level of "perceptual attributives" that function to attribute properties to particulars.

The doubly nonveridical character of pictorial experience can also be helpfully characterized using Susanna Schellenberg's (2013) recent account of discriminatory, perceptual capacities. To see an object with a particular 3D shape Σ, according to Schellenberg, is to exercise a capacity to differentiate Σ perceptually from other possible shapes and to single out particulars that instantiate Σ in the visible environment. When you hallucinate a Σ-shaped object, she maintains, the very same discriminatory capacity is employed. Your experience, however, fails to single out a particular in the environment and, for this reason, your employment of that capacity is baseless. Nonetheless, you are in a perceptual state whose content is of the same potentially particularizable type as the perceptual state formed when you are perceptually related in the appropriate way to a Σ-shaped object. In natural language, this shared content type would be articulated as "*That* object is Σ-shaped." (The content type is only potentially particularizable because, as in the case of hallucination, the demonstrative-like representational element may purport to single out an object in the visible environment without successfully doing so.)

Schellenberg's account can be usefully extended to the case of picture perception.[4] When you have an experience as of a Σ-shaped object in pictorial space, the very same discriminatory, selective capacity is triggered by the light reaching your eye as would be triggered by the light reflected from a Σ-shaped object seen face-to-face. Accordingly, although the experience elicited by the picture's surface is not successfully particularized, it contains the same type of content as the experience elicited when you are perceptually related to an object with that very 3D shape.[5]

---

4. I'm grateful to E. J. Green for this suggestion.

5. Matthen (2005) argues that ordinary language demonstratives and demonstrative-like, singular representational elements in visual experience do not apply to objects in pictorial space. This view is implausible. There are good reasons to suppose that the same mechanisms of perceptual segmentation, feature-binding, object-based selective attention, and other capacities for visual singular reference are engaged by sources of depth-specifying information in the light reflected from a pictorial surface as by sources of monocular and binocular depth information available when viewing a real-world scene. Indeed, empirical studies of these capacities in vision science, as emphasized in this treatment, typically exploit the methodology of picture-based or "virtual" psychophysics.
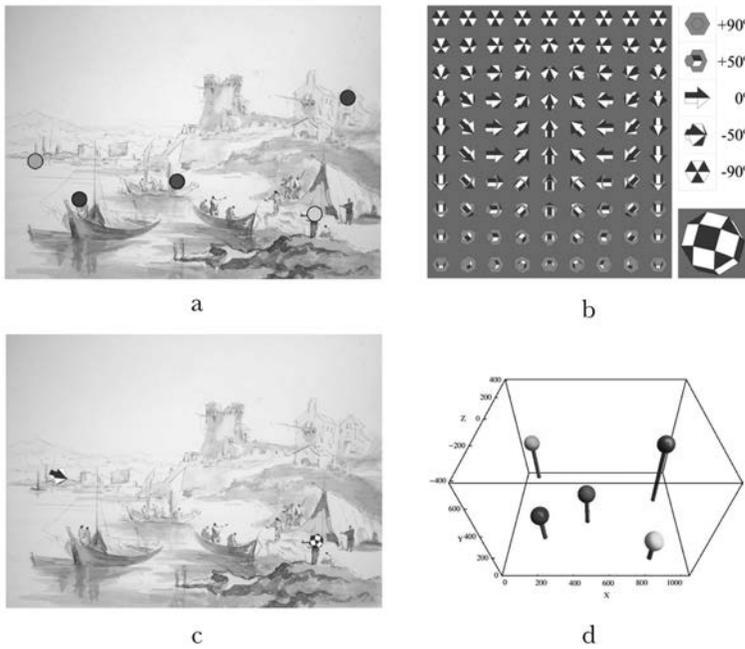
Figure 2. Measuring 3D point configurations in pictorial space

(a) A copy of a wash drawing by Francesco Guardi (1712–93), with 5 points selected for objects in foreground, middle ground, and background. (b) Different slant and tilt values for the pointer. (c) Subjects are instructed to modify the slant and tilt of the pointer in each trial using the arrow keys on a computer keyboard, until it points at the target. (d) The three-dimensional configuration computed for a complete set of pointer slant values. For 5 points, there are 5(5-1)=20 values. The colors of the points correspond to those in (a). From Wagemans *et al.* (2011). Measuring 3D point configurations in pictorial space. *i-Perception*, 2 (January 2011), 77–111, SAGE Publishing, doi:10.1068/i0420, http://ipe.sagepub.com/content/2/1. Images (a) and (c) reproduced with permission of Anne-Sophie Bonno (http://www.atelier-bonno.fr/galerie- copies-arts- graphiques.html).

Precisely which appearance properties are attributed by the visual system to objects in pictorial space can be ascertained using a variety of psychophysical measuring techniques (for reviews, see Koenderink et al. 2011 and Koenderink 2012). One method involves having subjects specify the attitude of numerous, small surface regions on an object experienced in pictorial space using a superimposed gauge figure (Koenderink and van Doorn 2003). Subjects are instructed to adjust the slant and tilt of a large number of such gauge figures for numerous, spatially distinct regions on the object. With this dense field of gauge figure settings, it is then possible to generate a depth map for the object or a computer-generated, 3D rendering of its relief in pictorial space. Another more recent technique explored

by Wagemans et al. 2011 involves specifying the perceived point-to-point directions for a number of point pairs in pictorial space (fig. 2a).[6] In each trial, a pointer and a target are superimposed on the image plane (fig. 2b). Subjects adjust the slant and tilt of the pointer in pictorial space until it appears to be oriented in the direction of the target (fig. 2c). For any given pair of points a and b in the configuration, two slant values are recorded, one for the direction $a{\rightarrow}b$ and one for the direction $b{\rightarrow}a$. Given n points, this process is repeated $n(n\text{-}1)$ times. Experimenters then use the slant values recorded for the pointer in a complete session to compute the best-fitting, 3D configuration (fig. 2d). Although there is considerable variability between subjects in the scaling of pictorial depth, 3D configurations for different subjects viewing the same picture have essentially the same shape. Like the gauge figure method, this technique enables vision scientists to operationalize the concept of depth in pictorial space.

Philosophical theories of depiction are typically concerned to explain, among other things, which aspects of pictorial representational systems distinguish them from nonpictorial representational systems, in particular, representation by means of language. One approach to dealing with this issue, common to experienced-based theories of depiction, starts from the idea that pictures make the properties that they depict their objects as having *visually present* to those who view them (Wollheim 1987, 1998, 2003; Hopkins 1998, 2006; Lopes 2005). Many philosophers, however, have assumed that empirical work in vision science isn't instructive when it comes to explaining the nature of pictorial experience, that is, when it comes to providing an account of the psychological, explanatory kind to which this experience belongs (an important exception is Gombrich 1961/2000). Vision science provides answers to "causal questions about the mechanics by which [pictorial experience] is generated" (Hopkins 1998, 113), but it doesn't throw light on the question concerning *what it is*, constitutively speaking, to have the experience as of an absent, 3D scene in a 2D surface.

I think this prevalent assumption is mistaken. In particular, both the deeply entrenched methodology of "virtual psychophysics" (Koenderink 1999) as well as empirical studies of pictorial space perception provide compelling support for the view that pictorial experience and seeing face-to-face are experiences of the same psychological, explanatory kind. In what follows, I shall refer to this view as the *Continuity Hypothesis*.

Almost all experimental studies of human and nonhuman vision in vision science rely on the methodology of virtual psychophysics: they confront subjects with photographs or computer-generated images of objects and scenes rather than their real-world counterparts. This is true of experiments on how subjects visually estimate environmental scene attributes such as distance in depth, orientation,

6. In physical space, this is the method of *exocentric pointing* (Cuijpers et al. 2000; Koenderink et al. 2008).

shape, size, texture, and lightness as well as studies of selective visual attention, object recognition, and organizational phenomena such as modal and amodal completion. It is also true, I should add, of most brain-imaging and electrophysiological studies of visual perception in neuropsychology. In "the study of vision," the psychologist James Cutting writes, "we must often look at an image of the world . . . ; the vicarious experience informs us about the direct experience" (2000, 635).

Such reliance on the methodology of virtual psychophysics in vision science is premised on the assumption that pictorial experience and seeing face-to-face are experiences of the same psychological kind. What justifies this assumption? The answer is that perceptual psychology individuates experiences, in part, by the sources of environmental information and information-processing mechanisms that produce them. "From the point of view of the methods of perceptual psychology," as Tyler Burge writes, "part of what it is to be a perceptual state of a certain kind is to be produced in something like the way that kind of state is produced by law-like formation principles" (2005, 20; see also Burge 2010, chap. 8). In particular, part of what it is to be a *visual perceptual state* is to be produced (1) in response to sources of optical information in the array of light reflected or emitted to the eyes from the environment and (2) in accordance with computational formation principles that are specific to the operations of the visual system.[7] The methodology of virtual psychophysics, as Burge writes in connection with experiments on apparent motion in pictorial displays, "takes advantage of the fact that a kind of perceptual state can be produced by a given [type of] proximal stimulation, whether or not the standard distal antecedents of the proximal stimulation are present" (Burge 2005, 23). To the extent that pictures and real-world scenes can cause the same types of information-bearing, proximal stimulation in their viewers, they can also cause their viewers to form perceptual states of the same psychological kind.

There is a substantial body of evidence that pictorial experience and seeing face-to-face are supported by the same bottom-up sources of optical information and produced in accordance with the same computational formation principles. Indeed, the various monocular sources of depth information used by the visual system to represent the spatial layout of a real-world scene are often referred to as "pictorial" depth cues precisely because they elicit corresponding experiences as of 3D spatial layout when present in the light reflected or emitted to the eye by a 2D, pictorial surface. These include, but are not limited to occlusion, texture gradients, shadows, reflections, relative size, familiar size, linear perspective, atmospheric haze, height in the visual field, and the horizon ratio (for surveys, see Cutting

---

7. These principles are often characterized in contemporary perceptual psychology within the framework provided by Bayesian causal inference theory. Precise characterization of them isn't necessary for present purposes, but for helpful, philosophically oriented discussions, see Clark 2013, Hohwy 2013, and Rescorla 2015.

and Vishton 1995; Cutting 1997; Palmer 1999, chap. 5; and Rogers 2003).[8] Some pictures, for example, stereograms, also take advantage of binocular sources of optical information such as retinal disparity, while movies make available potent, movement-based sources of distance and shape information such as motion parallax and the kinetic depth effect (Wallach and O'Connell 1953; Ullman 1979). "[P]erceiving depth in pictures and perceiving depth in the real world," as Cutting summarizing the empirical evidence concludes, "are cut from the same informational cloth" (2003, 236). This assessment, it should be emphasized, also extends to surface properties that are visually estimated, in part, on the *basis* of information specifying spatial layout (Fleming and Anderson 2004). The apparent lightness (albedo) of an object in pictorial space, for example, can be shifted from black to white by varying the perceived organization of the 3D scene in which it is located (Gilchrist 1977, 1980; Anderson and Winawer 2005).

Generalizing, we can formulate two Gibsonian principles:

> 1. For any determinate, spatial attribute *G*, e.g., a certain determinate shape, size, or orientation, a 2D surface *S* will elicit an experience as of *G* from a normally sighted viewer just in case the array of light reflected or emitted to the viewer's eye by *S* is of the same type as would elicit an experience as of *G* when reflected from or emitted by a real-world, 3D scene.

> 2. The light reflected from or emitted by *S* will be of this type only when it conveys one or more of the sources of optical information ordinarily used by the human visual system to represent the presence of *G* in a real-world scene. Many different token light arrays, it should be emphasized, can convey optical information about the same spatial attribute. For example, the arrays of light respectively reflected from a cubical wire framework seen face-to-face, a photograph of the framework, and a line-drawing of the framework can convey substantially the same information about 3D shape and orientation despite the numerous photometric differences between them (Hochberg 1962/2007, 1980/2007; Gibson 1971, 1979; Kennedy 1974).

Two points are important. First, psychophysical evidence that the spatial contents of pictorial experience and seeing face-to-face vary, in a closely corresponding way, with changes in the bottom-up, optical information available to the eyes indicates

---

8. Monocular depth cues support the experience of 3D shapes, orientations, relative sizes, and depth ratios in pictorial space, but they do not specify the distance of pictorial objects from the observer. As Vishwanath and Hibbard write, "under binocular viewing of pictures, although 3-D object shapes can be clearly perceived, their scale and absolute depth should remain optically unspecified" (2013, 1683). According to what Vishwanath 2014 calls the "absolute depth scaling hypothesis," this explains why most pictures fail to elicit a robust impression of *stereopsis* (solid appearance and immersive space) under binocular viewing conditions. Vishwanath emphasizes that the absence of stereopsis is not to be equated with the absence of apparent 3D structure: "*3-D structure* can be perceived solely on the basis of relative depth estimates but . . . the impression of *stereopsis* is induced only when absolute depth values can be estimated" (2014, 155).

that the visual system is processing that information—in both cases—in accordance with the same set of computational formation principles. It thus provides strong support for the view they are experiences of the same psychological kind (the Continuity Hypothesis).

Second, as Dominic Lopes (2005) has emphasized, perceiving the pictorial surface and its properties or "surface seeing" isn't psychologically necessary for pictorial experience.[9] A patterned surface, according to the principles above, will induce the experience of pictorial space just in case it reflects or emits light of a type that would elicit a 3D-scene-representing visual experience when reflected or emitted to the eye from the real world. (As Gibson puts it, "A picture is a surface so treated that a delimited optic array to a point of observation is made available that contains *the same kind of information* that is found in the ambient optic arrays of an ordinary environment" [1971, 31, emphasis added].) It isn't necessary that the surface *also* reflect or emit light of a type that would enable perception of the surface's own properties, e.g., its absolute distance from the perceiver's eyes, orientation, gloss, or texture. Viewing a single picture through a narrow aperture or two copies of the same picture through a stereoscope, for example, effectively eliminates cues that specify the presence and properties of the pictorial surface, while leaving accessible sources of optical information that specify depth and 3D structure in pictorial space (Ames 1925; Koenderink 2012; Vishwanath 2014). Other examples of pictures that divide pictorial experience from surface seeing include stereograms and autostereograms, 3D movies, large images seen at a distance, and successful instances of illusionistic painting (*trompe-l'oeil*). The experience of pictorial space doesn't depend on conscious (or even unconscious) perception of the properties instantiated by a 2D, pictorial surface.

In addition to sources of psychophysical evidence, there is also neuropsychological support for the Continuity Hypothesis: pictures of 3D objects activate the same processing areas in the visual brain as are activated by the depicted objects themselves. This is true not only of the object-categorizing ventral stream (see Grill-Spector et al. 2000 and Kanwisher 2004, e.g., for brain-imaging studies on how ventral stream areas respond to pictures of different kinds of objects), but also

---

9. This assessment seems, I might mention, curiously at odds with Lopes's own aspect recognition theory (ART) of depiction. According to ART, "pictures' designs present recognizable aspects of things . . . Pictorial recognition at this level may be called 'content recognition', since it consists in recognizing a *design* as the features making up an aspect of its subject" (1996, 145, my emphasis). A similar view is defended in a more recent treatment: "[P]ictorial recognition (unlike face-to-face recognition) is always the triggering by a two-dimensional surface of an ability to recognize a three-dimensional object. A flat surface triggers such a capacity only when the capacity to recognize the object has been extended so as to enable recognizing the object when it appears two-dimensionally. Thus pictorial recognition never occurs except by the mediation of a two-dimensional surface: *such a surface must be perceived*" (2006, 171, emphasis added). If, however, certain kinds of pictures divide pictorial experience—and, so, content recognition—from experience of the pictorial surface and its properties (Lopes 2005, chap. 1), then ART, as it stands, cannot be correct.

of the action-guiding, dorsal stream. Neurons in the intraparietal sulcus, a dorsal stream area involved in the skillful guidance of eye movements as well as reaching and grasping, for example, selectively respond to objects in pictorial space with different texture gradient and linear-perspective-defined 3D shapes and orientations (Taira 2001; James et al. 2002; Tsutsui et al. 2002, 2005; Sakata et al. 2003; Nelissen et al. 2009).[10] There is also evidence that neurons in area V3a are sensitive to depth in photographs (Berryhill and Olson 2009) and that neurons located in the lateral occipito-parietal junction, another early dorsal stream area, distinguish between images of graspable and nongraspable objects (Rice et al. 2007). Consistent with these findings, behavioral studies indicate the grasping system in the dorsal stream is able to exploit pictorial depth cues, although binocular cues are more heavily weighted, especially when grasping with the right hand (Marotta and Goodale 2001; Westwood et al. 2002; Gonzalez et al. 2008). Westwood and colleagues go so far as to suggest that the dorsal grasping control system "does not appear to discriminate in a fundamental way" between objects seen face-to-face and their counterparts in pictorial space (2002, 267). The ability to distinguish between objects and pictures of objects, they suggest, is rather a function of high-level, object recognitional systems in the ventral processing stream.[11]

There are a variety of phenomenologically salient respects, of course, in which the 3D-scene-representing experience elicited by a picture is normally distinguishable from that of seeing a 3D scene face-to-face (genuine cases of pictorial deception or *trompe l'oeil* are rare and obtain only under special viewing conditions). First and foremost, in some cases of pictorial experience, we are visually aware of at least some of the properties exemplified by the 2D, pictorial surface. It may not be psychologically possible to experience the same solid angle in the visual field as filled by a nontransparent, 2D surface at a single distance in depth and by an array of 3D objects located at different distances in depth simultaneously (Gombrich 1961/2000; Newall 2015; Briscoe forthcoming). But plausibly it is possible, when looking at a picture, to divide attention between the region of

---

10. There is conflicting evidence concerning whether 3D shape from shading, a potent source of spatial information in pictures, is processed in the dorsal stream. For discussion, see Georgieva et al. 2008.

11. Several philosophers including Mohan Matthen (2005, chap. 13) and Bence Nanay (2010, 2014) have proposed that objects in pictorial space are represented by the object-categorizing, ventral processing stream, but are not typically represented by the action-guiding, dorsal processing stream. The evidence reviewed here speaks against this proposal: there are empirically good reasons to believe that the dorsal stream is far from blind to (virtual) depth and 3D structure in pictorial space. That said, there are good reasons to think that the objects of pictorial experience are not represented by the dorsal stream in the *same way* as the objects of ordinary, nonpictorial visual perception. One crucial difference is that sources of absolutely scaled distance information required to program certain motor actions are not typically available for the objects of pictorial experience (Briscoe forthcoming; Cohen and Meskin 2004; Vishwanath 2014; see note 9). This would explain why objects on the "far side" of the pictorial surface as Matthen and Nanay observe, do not normally appear to be potentially touchable.

phenomenally 3D, pictorial space contained within some solid visual angle θ and the distribution of 2D, pictorial surface properties contained within some different solid visual angle φ. Similarly, it is possible to experience textual elements visible on a pictorial surface as in front of and partially occluding objects presented in pictorial space. A word printed on a pictorial surface may be visually experienced as a figure on a more distant, three-dimensionally organized background, and, when this is the case, we experience properties of the real-world surface and virtual scene at the same time. Second, experiences of pictorial space convey rich information about relative distance in depth (object a in pictorial space, e.g., may be experienced as about twice as close to the pictorial point of view as object b), but they do not typically represent their objects as located at certain absolutely scaled distances in depth (Vishwanath 2014; Briscoe forthcoming). Pictorial experience is highly indeterminate with respect to the egocentric, viewer-relative locations of its objects in a way that ordinary visual experience is not. In consequence, it does not typically give rise to a robust impression of *stereopsis*, i.e., solid appearance and immersive space. Third, the geometry of pictorial space is "highly volatile" in comparison with the geometry of visually perceived, physical space (Koenderink 1997): perceptual representations of depth and 3D shape in pictorial space vary significantly across changes in viewing angle, distance, and, in the case of photographs, lens size. Finally, when a subject moves in relation to a stationary, real-world object, the object's orientation doesn't visually appear to change. By contrast, when the subject moves in relation to certain pictures, for example, the famous British army recruiting poster depicting Lord Kitchener (Gombrich 1972, 113), the object she experiences in pictorial space may curiously appear to rotate toward her. (Discussion of other phenomenologically salient contrasts between pictorial experience and seeing face-to-face can be found in Thompson et al. 2011, chap. 12.)

Although space doesn't permit developed discussion, these differences, I would suggest, do not motivate skepticism about the Continuity Hypothesis. On the contrary, phenomenologically distinctive features of pictorial experience are explained by vision scientists using the same theoretical framework as is used to explain seeing face-to-face. Central to relevant explanations are a specification of the types of property-specifying information present (or absent) in the light reflected or emitted by a pictorial surface under different sets of viewing conditions as well as a description of the computational formation principles by which the operations of the visual system are normally governed.[12] This is to say that, from the standpoint of mainstream vision science, each of the four differences above is to be conceived as a difference at the level

---

12. These principles will determine, among other things, how available sources of information are to be weighted and integrated as well as which sources of information specify properties of objects in pictorial space and which properties of the 2D, pictorial surface.

of *visual representational content* and, so, as a difference between experiences of the same psychological kind.[13]

The task of a philosophical account of pictorial experience, Hopkins writes,

> is to characterize that experience and to use that characterization to provide an analysis of the notion of depiction. It is no part of that account to say *why* we see one thing rather than another in a surface. That question does not fall to the philosopher at all, but to the art historian, the psychologist or the physiologist. It concerns the causal framework out of which the characterized experience is born. It is important that what we as philosophers say is consistent with whatever turns out to be the answer to this empirical question. What is not necessary is that philosophy itself provide the answer. (1998, 112–13)

Vision science, I have argued here, doesn't merely illuminate the causal framework out of which pictorial experience is born. It also throws light on the *nature* of that experience: on what it is, constitutively speaking, to have the experience as of depth and 3D structure when looking at a 2D surface. In particular, empirical studies of pictorial experience as well as the methodology of virtual psychophysics provide strong support for the view that pictorial experience and seeing face-to-face are experiences of the same psychological, explanatory kind.

2.2. DEPICTION AS VIRTUAL MODELING IN PICTORIAL SPACE

Richard Wollheim was famously skeptical about the possibility of providing a general account of the conditions under which a 2D surface will elicit a 3D-scene-representing visual experience:

> [Pictorial experience] is triggered off by the presence within the field of vision of a differentiated surface. Not all differentiated surfaces will have this effect, but I doubt that anything significant can be said about exactly what a surface must be like for it to have this effect. (Wollheim 1988, 46)

Empirical studies of pictorial experience and the methodology of virtual psychophysics in vision science provide reason to think that such skepticism is unfounded. In fact, they suggest that just as much can be said about what the light reflected from a 2D surface must be like to elicit an experience as of a certain 3D spatial layout as can be said about what the light reflected from the world must be like in order to have this effect. While vision science has a lot to tell us about the

---

13. For discussion of "conjoint" representation of pictorial space and the pictorial surface, see Mausfeld 2003, Millar 2006, and Briscoe forthcoming; for discussion of the absence of absolutely scaled depth information and stereopsis in pictorial experience, see Koenderink et al. 1994, Vishwanath and Hibbard 2013, and Vishwanath 2014; for discussion of the effects of viewing angle, viewing distance, and lens size, see Cutting 2003, Koenderink and van Doorn 2003, Sedgwick 2003, and Thompson et al. 2011, chap. 12; for discussion of apparent object rotation in pictures, see Goldstein 1987, Busey et al. 1990, and Koenderink et al. 2004.

nature of pictorial experience, it doesn't, however, equip us with an account of *depiction*. Indeed, the psychological explanation of pictorial experience characterized above isn't specific to pictorial artifacts. It also applies to the "accidental images" discussed by Wollheim in *Painting as an Art* (1987, 46–59). The deep resemblance theory, presented in this section, begins where work in vision science leaves off.

Pictorial space, it seems clear, is a richly productive arena for the construction of *virtual models*. If you are an architect and you want to convey information about the visual appearance of a house, you could construct a precise scale-model out of balsawood or some other material. Alternatively, you could construct a virtual, 3D model of the house using computer-aided design (CAD) software. For many purposes, a virtual pictorial model is preferable to a real-world model: it can be stored and transmitted electronically, and it can be modified on demand in an open-ended number of ways. (Don't like the aspect ratio of the windows or the color of exterior walls? Want to take out the third bathroom and add a closet? No problem!) Similarly, if you are vision scientist and you intend to do psychophysical research on how the mechanisms of lightness constancy and depth perception interact, you could construct an apparatus in which the distance, orientation, and illumination of a physical target stimulus can be independently varied. Alternatively, using stimulus presentation software, you could construct a virtual scene in pictorial space and modify stimulus parameters for a virtual target instead. For many purposes (but not all, see Koenderink 1999), virtual psychophysical objects and scenes are preferable to their real-world counterparts. Among other things, they are less time-consuming and expensive to construct, stimulus parameters can be modified with ease, and they can be designed to work with customized response collection tools.

These examples bring me to my core claims. According to the deep resemblance theory, a marked or otherwise patterned surface S is a picture just in case two conditions are met. First, S must present its viewers with a *virtual model* of an object or 3D scene. *The Hunters in the Snow* (fig. 1), for example, depicts, among other things, hunters and dogs because its surface elicits, by design, an experience as of voluminous objects in pictorial space, which objects were intended by Bruegel visually to resemble hunters and dogs in certain salient respects. I take Leonardo to have had something like this view of the relation between depiction and pictorial experience in mind when he wrote, "The first intention of the painter is to make a flat surface display a body as if modeled and separate from this plane" (quoted in Kemp 1989, 15).

The first condition does not specify *how* a surface must achieve this end in order to qualify as a picture. It leaves open the possibility of what Roberto Casati refers to as a "hallucinatory picture," that is, a surface that reliably induces a visual-hallucination-like experience as of an object or 3D scene, but in a way that cannot be psychologically explained by the properties visibly instantiated on the surface (Casati 2010). "Hallucinatory pictures," Casati writes, "are accompanied by a sense of magic, of unexplained causation" (2010, 366).

The second condition fills this explanatory gap. In order to qualify as a picture, not only must a surface S elicit the experience as of an object with certain properties in pictorial space, the content of this experience must be systematically guided by sources of optical information in the light reflected (or emitted) by S to the eye. For example, if S is designed to elicit the experience as of a Σ-shaped object in pictorial space, then S must reflect light of the same type as would elicit experience as of a Σ-shaped object when reflected from a real-world scene. Moreover, relevant sources of optical information in the light reflected from S and transduced by the retina must be processed in accordance with the same law-like formation principles that govern seeing face-to-face (section 2.1). This second condition rules out the possibility of a hallucinatory picture. A surface S is a picture only if looking at it causes the experience of virtual depth and 3D structure in the right way.[14]

According to the deep resemblance theory, then, a surface S functions as a picture only if S presents its viewers with a virtual model of an object or scene rendered in phenomenally 3D, pictorial space. And a surface S depicts an object O only if the virtual model presented by S is a model of O, that is, just in case O is the model's original. The deep resemblance theory thus rejects the assumption that depiction-relevant resemblances between a picture and its subject are superficial—that the model a picture presents of its original is visually experienced as spatially flat or 2D. According to the theory, we need to look into pictorial space, *through* the pictorial surface, as it were, to experience the R-properties in virtue of which a picture depicts its subject. R-properties need not be limited to those that can be instantiated by a 2D design or pattern. Instead, they encompass these as well as other properties that objects may be seen to instance when encountered face-to-face. Importantly, these include not only intrinsic properties such as 3D shape and lightness, but also nonintrinsic, viewpoint-dependent properties. The depth ratio $(d_1:d_2)$ between two objects in pictorial space and their respective directions from the pictorial point of view, for example, may be intended to resemble the depth ratio between two objects in physical space and their respective directions when viewed from a certain real-world location.

Other recent resemblance theories restrict R-properties to those that abstract from the dimension of depth (Budd 1992/2008, 1993/2008; Hopkins 1998, 2006; Abell 2009). They look to visual appearance properties that can be instantiated by both a picture's 2D surface and the object it depicts. The deep resemblance theory, by contrast, doesn't single out any set of intended resemblances as privileged. A virtual model may be intended to resemble its original O in respect of any combination of visual appearance properties that O may be seen to instance when encountered face-to-face. The deep resemblance theory paints with a rainbow palette of R-properties.

---

14. I'm grateful to Alberto Voltolini for prompting me to be clearer on this point.

Such nonrestrictivism about R-properties makes good sense for two reasons. First, real-world models are evidently intended to resemble their originals in an open-ended variety of ways. Which correspondences or isomorphisms are set up between a real-world model and its original depends on the way the model is *used*, on the specific work that is done with the model, on which questions, if any, it is supposed to answer. The same is true of a virtual model constructed in the "medium" of pictorial space: the respects of resemblance that matter depend on the activity or inquiry in the service of which the model is constructed.

Second, nonrestrictivism about R-properties enables the theory to explain depiction in a way that permits it to meet what Dominic Lopes calls the "diversity constraint" (1996, 32). An adequate resemblance theory of depiction not only must explain *what it is* for a surface to function as a picture, it must do so in a way that respects the diverse range of pictorial styles and types. In other words, it must explain depiction in terms of resemblances that are not confined to any particular method or set of conventions for producing pictures, for example, the method of axonometric projection or Egyptian profile style.

The problem, Lopes argues, is that while intention-based resemblance theories have the resources to "reconcile resemblance with the diversity of depiction" (1996, 19), they also run the risk of circularity (also see Goodman 1976, 39). In particular, they run the risk of appealing to *depiction-dependent* resemblances, resemblances that can only be discerned given prior knowledge of which object is being depicted. "[P]ictures in different styles share a number of disjoint sets of properties with their subjects. The capacity to see how an *ukiyo-e* print, a Byzantine icon, a cubist collage, or a Kwakiutl totem each resembles its subject seems to depend at least in part on a familiarity with how each kind of picture [depicts]" (Lopes 1996, 19). Our ability to discern intention-based resemblances between a picture *P* and an object *O*, Lopes suggests, is a consequence of our ability to identify *O* as the object depicted by *P* rather than the other way around.

My reply to this objection is threefold. To begin with, Lopes's objection, as far as I can tell, is purely conjectural. Claims about the limits of human abilities to perceive or otherwise discern similarities between the objects they experience, however, should be empirically motivated: they should take account of psychological evidence concerning the different types of similarities that human beings are actually capable of picking out and the various structure-mapping abilities that enable their detection (Medin et al. 1993; Gentner and Markman 1997; Goldstone 1994; Ramscar and Hahn 2001). Lopes, however, presents no psychological evidence that intention-based visual resemblances are, in general, depiction-dependent.

Second, the objection fails to give intention-based resemblance theories the benefit of two assumptions that Lopes allows his own aspect recognition theory of depiction. One is the assumption that human capacities for visual object recognition are *dynamic*. We are reliably able, e.g., to identify an object as a horse, or a house, or a boat, across significant variation in the object's appearance and viewing conditions. As Lopes puts it, "objects recognized are not experienced as similar

to objects previously seen in uniform ways" (Lopes 1996, 222). The other is the assumption that a picture's subject need only be identifiable by suitable perceivers, where suitable perceivers not only have the capacity to recognize the picture's subject in the flesh, but are, moreover, competent in the style to which the picture belongs. "Somebody who can recognize Vollard but cannot interpret Cubist pictures is not a suitable perceiver of [Picasso's painting of Vollard]" (Lopes 1996, 153). The proponent of intention-based resemblance can advert to both of these assumptions. Just as objects recognized in the flesh are not experienced as similar to previously seen objects in uniform ways, pictures in different styles need not be experienced as similar to their *depicta* in uniform ways. Moreover, explanatorily relevant similarities between a picture and its subject need only be recognizable by perceivers already competent in the style or system of depiction to which the picture belongs.

Last, the objection clashes with currently influential models of visual object recognition in cognitive science. In particular, influential prototype and exemplar theories take visual object recognition to be, at core, a *similarity-based* process (for overviews, see Murphy 2002 and Machery 2009). According to prototype theories, for example, ascertaining whether a perceived object O belongs to a certain category C involves generating an internal representation of O's visible properties; retrieving from long-term memory a representation of the visible properties statistically associated with C's membership; computing the degree of similarity between these two representations; and, last, applying a decision rule that specifies the degree of similarity required for membership in C. Exemplar theories primarily differ from prototype theories in treating stored object representations involved in recognition as representations of previously encountered category members. That said, the categorization process is no less similarity based. Identifying an animal as a cat, e.g., involves computing the degree of similarity between the animal's perceived properties and those of previously encountered cats. On both approaches, visual resemblance is grist for recognition's mill. Hence, if Lopes is correct to suppose that "pictorial recognition involves some of the same processing as implements recognizing objects face-to-face" (Lopes 2006, 170), then there is good empirical reason to think that the former is a similarity-based psychological process. And since salient visual resemblances between a picture and its object—those that can presumed to drive the process of pictorial recognition—are typically intended (nonaccidental), it follows that our capacity to detect intention-based resemblances is not, in general, psychologically dependent on prior identification of which object the picture depicts.

2.3. AGAINST THE CONTENT THESIS

Almost all philosophical discussions of depiction in the last fifty years have been structured by the assumption that a theory of depiction is equivalent to a theory of pictorial representation. "Philosophical studies of depiction," Katerina Bantinaki writes, "focus on the representational function of figurative pictures: they aim

to explain how such pictures represent and how pictorial representation relates to other types of representation" (2009, 238). The deep resemblance theory breaks with the assumption that depiction is a genus of representation (the Content Thesis). I argued earlier that only some real-world models have an iconic representational function. Many models rather perform a function that I described as broadly *substitutive*. The same is true, I shall now argue, of virtual, pictorial models: only some pictorial models—and, so, the pictures that present them—are in the representation business. The task of providing a theory of depiction is thus prior to the task of working out a theory of how pictures are used to represent the visible world.

The claim that at least some pictures present viewers with nonrepresentational substitutes or surrogates isn't entirely original. A well-known version of it can be found in E. H. Gombrich's "Meditations on a Hobbyhorse" (1951/1985). The deep resemblance theory goes beyond Gombrich's "substitute account" in two important ways, however. First, the deep resemblance theory maintains that substitutive pictorial models are not experienced as on a picture's surface, but as *in* its surface. Pictorial models are *virtual models* constructed in phenomenally 3D, pictorial space. The properties in virtue of which they intentionally resemble depicted objects, accordingly, include many of the 3D properties attributable to the objects of ordinary, nonpictorial visual experience.

Second, the deep resemblance theory provides criteria for distinguishing substitutive pictorial models from those that perform an iconic representational function. More specifically, a pictorial model is an *icon* only when it is used in ways that meet analogues of the Guidance and Systematicity requirements outlined earlier:

> *Guidance (for pictures)*
> In order iconically to represent its original *O*, a pictorial model *M* must have the function of guiding the way its viewers perform some task *T*, where *T* involves either engaging in actions directed in relation to *O* or forming perceptually based beliefs about *O*'s visible attributes, and, further, it is a condition of the (nonaccidental) successful performance of *T* that *M* actually resemble *O* in intended, visible respects.

> *Systematicity (for pictures)*
> The way viewers are guided by a pictorial model *M* in performing *T* must systematically depend on *M*'s visible attributes, such that had the intention-based respects in which *M* visually resembles *O* been different, then viewers' performance of *T* would have proved successful only if *O*'s visible attributes had been different in corresponding respects.

It follows from these requirements that depiction is necessary, but not sufficient for iconic pictorial representation. That a given virtual model iconically represents an object or scene is a contingent, empirical fact about how the model is used, and not something constitutive of depiction in general. Just as models can be constructed in the real world for various nonrepresentational purposes, models can be constructed in pictorial space for various nonrepresentational purposes. There is nothing special about the medium of pictorial space that automatically confers a representational status on a model constructed within it. Here are two examples to illustrate this point.

Consider, first, *trompe-l'oeil* paintings, in which pictorial experience is "divided," as Lopes 2005 puts it, from seeing a picture's design properties. Wollheim argues that *trompe-l'oeil* paintings are not representational because they are not intended to be experienced as pictures: "they do not invoke, [but rather] repel, attention to the marked surface" (Wollheim 1987, 62). Given the commitments of Wollheim's "seeing-in" theory of depiction—a surface is a picture only if it elicits, by design, an experience that represents among other things the markings on that surface—it follows that *trompe-l'oeil* paintings do not, strictly speaking, depict their subjects. As paradoxical as it may sound, a *trompe-l'oeil* painting of an object or scene is not, strictly speaking, a picture of that object or scene.

I think we should accept the conclusion that *trompe-l'oeil* paintings do not typically have a representational, property-attributing function without inferring, as does Wollheim, that they are not pictures. Depiction, on the view defended here, is necessary, but not sufficient for pictorial representation. Some pictures present virtual models that perform a substitutive rather than a representational function. *Trompe-l'oeil* paintings of domes, apses, and other architectural elements are prime examples of this. A *trompe-l'oeil* painting of a pilaster, for example, doesn't have the function of attributing certain properties to a pilaster. Rather its function is to sustain the illusion that a pilaster with certain properties is "substantially present" (Feagin 1998)—that is, to produce some of the perceptual effects of seeing a real pilaster in its viewers.

Consider, second, the widespread use of virtual, pictorial objects and scenes in psychophysical and neuroscientific investigations of vision, discussed earlier (section 2.1). The photographic and computer-generated images in which these virtual stimuli are experienced, it seems clear, do not have a representational, property-attributing function. Their purpose is not to *inform* experimental subjects about the visible properties of some actual or possible object or scene. Rather, the function of a virtual stimulus, in context, is to go proxy for a real-world stimulus with certain visual attributes. Again, the point of the example is that virtual pictorial models, like real-world models, need not perform a communicative, representational function for those who make or view them. The Content Thesis is false.

2.4. AGAINST THE VEHICLE THESIS

I now turn now to the Vehicle Thesis. According to the Vehicle Thesis, the vehicle of a picture's representational content *C* is the design visible on the picture's surface. In this respect, if not others, a picture that represents an object as having certain properties is like a written description of the object. The vehicle of representational content, in both cases, is a marked or otherwise differentiated, 2D surface. Commitment to the Vehicle Thesis is explicit in an influential treatment by Robert Hopkins:

> We understand pictures by looking at them, but this is equally true of the expressions of written language. In both cases we are aware that we are looking at a representation, and in both our experience presents us with marks on a surface, thus differing quite sharply from the experience

of seeing the represented object in the flesh. In what, then, does the visual nature of picturing, as opposed to language, lie? (1998, 3)

If, as Hopkins elsewhere writes, "[w]hen we see a marked surface, what our experience presents us with is *just a set of marks*" (2006, 147, emphasis added), then how is seeing and understanding a picture of an object, in general, different from seeing and understanding a written description of that object?

Hopkins's answer is that pictorial understanding depends on an experience of a certain kind of resemblance between a picture's design and its subject. According to his experienced resemblance in outline shape theory (EROS), in order to see and understand a picture as depicting a horse with certain properties, for example, it is necessary (1) to see a region R on the picture's surface and, further, (2) to experience R as subtending the same solid visual angle as a horse or, as Hopkins puts it, as resembling a horse in "outline-shape" (Hopkins 1998, 2006). The vehicles of pictorial representation, on this view, are thus surfaces that are experienced as having a certain 2D spatial organization.

The deep resemblance theory rejects the Vehicle Thesis. According to the former, the proper vehicles of pictorial representational content are virtual models. Virtual models, unlike written descriptions, are not real-world entities. They do not exist independently of our perceptions of them—their *esse* is *percipi*. The reflexively representing properties of a pictorial model *qua* icon are not experienced as on the mind-independent, 2D pictorial surface, but rather as in phenomenally 3D, pictorial space.[15]

There are two main reasons to be skeptical of the Vehicle Thesis, I would suggest. First, seeing the pattern on the 2D pictorial surface isn't psychologically necessary for pictorial experience. Some pictures, as pointed out above (section 2.1), sustain pictorial experience in the absence of conscious or unconscious representation of pictorial surface properties. In consequence, the putative analogy with textual representation breaks down.

Reading text involves forming perceptual (visual or tactile) representations of marks made on an external surface. Indeed, those external marks are *words*—vehicles of textual representational content—only in virtue of the functional role that representations of them play in processes of linguistic comprehension and production. By contrast, in order to exploit the information conveyed by a representational picture it isn't necessary to perceive its 2D design. It is possible, for instance, to see and make sense of a photograph or computer-generated image that represents a house as having certain properties in the absence of optical information that specifies properties of the picture's surface—say, because you are viewing

---

15. Wiesing 2009 similarly distinguishes between (a) the pictorial surface (the *image-carrier*); (b) the "exclusively visible" object experienced in pictorial space (the *image-object*); and (c) the depicted object (the *image-subject*), which (b) is intended visually to resemble. For Wiesing, (a) displays but does not represent (b). When a picture is used for representational purposes, it is (b) that functions to attribute properties to (c).

the picture through a narrow, monocular aperture (Ames 1925; Schlosberg 1941; Vishwanath and Hibbard 2013.) Representations of the picture's design properties in this case are not formed by the visual system and, so, clearly are not *used* by it for purposes of working out what the picture depicts. From an information-processing standpoint, making sense of a picture of a house can utilize the same perceptual and cognitive mechanisms as are utilized in making sense of a 3D, real-world model of the house (when seen from a certain stationary point of view). This suggests that a picture's 2D design isn't the vehicle of its representational content. However, it is consistent with the claim that it is an object that we experience in the picture—a virtual model, to use my terminology—that is doing the work of representing.

Second, the Vehicle Thesis assumes that our ability to work out a picture's representational content is explanatorily dependent on seeing the picture's design, on visually experiencing the structure or composition of a 2D pattern on the pictorial surface. Variations in the experienced structure or composition of the picture's design are supposed to explain differences in how we interpret or make sense of the picture. The problem, however, isn't just that it is possible successfully to interpret a picture without experiencing the picture's design. The problem is that when we *do* experience a picture's design, the properties attributed by our visual system to the design typically depend on the properties attributed by our visual system to virtual objects in pictorial space. In particular, the apparent 2D spatial organization of the design on a picture's surface is typically dependent on the apparent 3D spatial organization of the scene that we experience in the picture's surface.

This point can be brought out by reflecting on Hopkins's experienced resemblance in outline shape (EROS) account of depiction. According to Hopkins, to experience a certain 3D scene in a picture's surface it is necessary to experience the 2D outline shapes of certain regions on the surface as resembling those of the objects in that scene. In what follows, I shall refer to the perceived 2D, spatial organization of the pictorial surface determined by these visually represented outline shapes as the "outline configuration" of the picture's design (or its *O-configuration* for short).

According to EROS, seeing the *O*-configuration of a picture's design, is explanatorily prior to pictorial understanding. It is only because we experience the outline shapes visible on the surface of *The Hunters in the Snow* (fig. 1), for example, as resembling the outline shapes of objects in a certain three-dimensionally organized scene that we are able to make sense of the former as depicting the latter. EROS, however, takes our visual awareness of the *O*-configuration of a picture's design for granted: it provides no explanation of how viewers on the basis of information available to the eye perceptually segment regions on the picture's surface that correspond to the outline shapes of objects in the depicted, 3D scene.

The problem is that there is nothing intrinsic to a picture's photometric configuration—that is, to the way light-reflecting marks are distributed across its surface—that dictates a particular *O*-configuration (Briscoe 2008). On the contrary, in most, if not all cases, it is the apparent, 3D organization of the scene we

experience *in* the picture's surface—in pictorial space—that determines which array of 2D outline shapes we experience as instantiated *on* the picture's surface. In other words, the picture's perceived *O*-configuration is determined by the way in which regions of the scene experienced in pictorial space are segmented and grouped together in depth.[16]

This point can be developed by reflecting further on the role of perceptual organization in pictorial experience. Consider the variant of Edgar Rubin's Maltese Cross presented in figure 3. The visual experience evoked by the Maltese Cross is ambiguous or multi-stable, meaning that, with prolonged viewing, figure-ground assignments can alternate. Importantly, it isn't only the relative depth relations that "flip" between these assignments. There are also changes at the level of the objects and shapes that are visually represented in our experience. In assignment (a), an upright, gray cross appears in front of a white, amodally completed square; in (b), a white propeller-shape appears in front of a gray, amodally completed diamond; while, in (c), we experience a modally completed gray and white diamond in front of a white, amodally completed square.
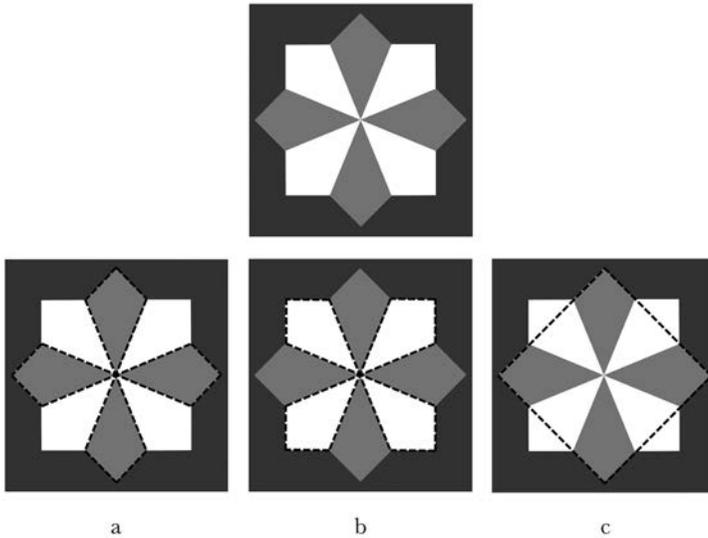

Figure 3. Rubin's *Maltese Cross*

The visual experience evoked by the *Maltese Cross* is ambiguous (multi-stable).
In figure-ground assignment (a), an upright, gray cross appears in front of a white,
amodally completed square; in (b), a white propeller-shape appears in front of a gray,
amodally completed diamond; and in (c), we experience a modally completed gray
and white diamond in front of a white, amodally completed square.

16. I am grateful to Robert Hopkins for discussion of this claim.

Importantly, to each of these different figure-ground assignments, there corresponds a different *O*-configuration on the picture's surface. There is nothing intrinsic to the photometric configuration of the Maltese Cross that determines *which O*-configuration we perceive. Rather, the perceived *O*-configuration varies with the apparent depth ordering of the scene we represent when we look at the image, and the latter varies in turn with how innate or learned mechanisms of visual perceptual organization group and segment surface regions in phenomenally 3D, pictorial space (while the photometric configuration remains the same).

This second example is a simple one. But it clearly illustrates the point that the way we experience the spatial properties of the design visible *on* a 2D, pictorial surface is significantly shaped by our experience of the organization of the 3D scene we experience *in* the surface. Which outline shapes are experienced as making up a picture's design crucially depends, in part, on the spatial layout of objects and surfaces we experience in phenomenally 3D, pictorial space. The photometric configuration of a picture's surface no more by itself determines its perceived *O*-configuration than the photometric configuration of the retinal image by itself determines the perceived spatial configuration of the objects present in the scene before our eyes. (Hence, the so-called inverse optics problem in perceptual psychology.)

What holds for the stimulus presented in figure 3, it should be emphasized, holds equally for figure 1 and for other pictures in which we experience a three-dimensionally organized scene of some kind: figure-ground assignments in phenomenally 3D, pictorial space dictate which 2D outline shapes we experience as instantiated on the pictorial surface. But, if this is right, then, contrary to EROS and the Vehicle Thesis, a picture's perceived 2D design doesn't stand in the right psychological relations to psychological processes involved in working out the picture's content to function as the latter's vehicle. Representing the properties and organization of objects in pictorial space has psychological priority relative to representing the properties and organization of the design on the pictorial surface.

What, then, is the correct characterization of the role of 2D, pictorial design properties in making possible pictorial representation and pictorial understanding possible on this revisionary account? The answer, I would suggest, is that pictorial design properties play a causal, enabling role. When all goes well, the sheaf of light rays they reflect or emit triggers an experience as of an object with certain visible attributes in phenomenally 3D, pictorial space, which visual attributes sometimes reflexively represent those of a real-world object.[17] It is a task for perceptual psychology rather than philosophy to explain in detail the different conditions under which this triggering will take place, a task on which, as indicated above, substantial progress has already been made.

---

17. I would emphasize that it is a merely contingent fact about historical, picture-making practices that artists and others have standardly employed technologies that, like written communication, involve marking or otherwise physically altering the light-reflecting properties of a 2D surface. This is true of drawings, paintings, etchings, and photographs, but it isn't true of computer-generated images.

3. SOME CONSEQUENCES OF THE
DEEP RESEMBLANCE THEORY

The deep resemblance theory breaks with a dominant approach to depiction structured by Content and Vehicle Theses. It offers in the later Wittgenstein's sense a new "picture" (*Bild*) of depiction. A picture in the relevant sense, Gordon Baker writes,

> picks out certain things as self-explanatory, others as problematic; it steers attention towards certain aspects of things and away from others; and it guides the direction that problem-solving takes and helps to set the standard of adequacy for a solution. . . . [A picture] determines a whole intellectual orientation. (2001, 13)

In this concluding section, I would like briefly to lay out some of the more interesting consequences of the intellectual orientation toward depiction and pictorial experience advocated here.

First, the thesis that pictures are *transparent* to their subjects (Walton 1984, 2008; Lopes 1996, 2003) is false. The proper object of pictorial experience isn't the depicted object itself, but a virtual model of the depicted object rendered in pictorial space. We no more see the Stonborough House itself when we look at a drawing or computer-generated image of the house Wittgenstein designed for his sister Margaret than when we see a real-world, architectural model of the house face-to-face. Similarly, we no more see a horse when we look at a painting or an etching of a horse than, say, when we look at an equestrian bronze or a horse made of stone.

The deep resemblance theory is clearly inconsistent with the transparency thesis, but are there any independent reasons to think the transparency thesis is false? From the standpoint of the deep resemblance theory, there had better be! After all, if the proper object of pictorial experience were the depicted object itself, then, as Robert Hopkins (in personal communication) writes, "once we've accounted for what [pictorial experience] is, we've done all the work that needs doing. No work for an appeal to deep resemblance remains."[18] If, for example, we really do see a *horse* when we look at George Stubbs's life-size painting of Whistlejacket, then there would be no point in appealing to intended resemblances between the object we experience in the pictorial surface and a horse for purposes of explaining how the painting depicts (see Wollheim 2003, 140). So the question, now, to ask is: How plausible is the transparency thesis? If, by hypothesis, we don't see an object properly categorized as a horse when we look at a 3D sculpture of a horse, then why ought we to think that we *do* see such an object when we look at an equestrian painting?

---

18. It is important to emphasize the conditional nature of this claim: Hopkins doesn't actually accept the transparency thesis (2012, 713).

Dominic Lopes has argued that viewers are able to see through a picture to the depicted object "because the conditions under which pictures function parallel the conditions under which we experience the objects of visual perception" (1996, 192). In particular, the following three conditions jointly obtain:

1. The depicted object O is a cause of the design D visible on the picture's surface.

2. D is counterfactually dependent on O. Had O's visible properties been different, then D's properties would have been different too.

3. There is second-order isomorphism between O's visible properties and D's visible properties: similarities among the former mirror similarities among the latter.

Some philosophers, including Kendall Walton (1984, 2008), have argued that in addition to (1)–(3), a picture must also be mechanically executed in order to support transparent experience. I won't rehearse Walton's familiar reasons for this restriction here. I will, however, assume that if mechanically executed pictures such as photographs are not transparent in virtue of meeting conditions (1)–(3) above, then handmade pictures are not transparent either.

Are conditions (1)–(3) sufficient for photographic transparency? Here is an example that suggests they are not. 3D printers use a computer-guided, additive, layering process to replicate an object's volumetric structure with exquisite accuracy. The real-world models that result from this mechanical process faithfully meet all of the conditions that Lopes and Walton require for pictorial transparency. It seems clearly false, however, that in seeing a model of an object O produced by a 3D printer we are literally seeing O itself as opposed to something that non-accidentally instantiates many of O's visible properties. But, if this is so, then, by analogy, there is no reason to suppose that in seeing a photograph of O that meets conditions (1)–(3) we are literally seeing O itself. And, if transparency fails for photographs, then it also fails for handmade pictures such as drawings, paintings, and etchings. Pictorial experience is, in general, opaque.

I've argued that pictures aren't transparent to the individuals that they depict. Both real-world and virtual pictorial models, however, are sometimes transparent to the *properties* of their originals. Visually experiencing a model of the Stonborough House, whether face-to-face or in pictorial space, can be a way of vicariously experiencing the shapes, colors, and other visual appearance properties instantiated by the building itself. This is a direct consequence of the reflexivity that is characteristic of iconic representation (section 1). It captures, I take it, the heart of the intuition that representation by means of picturing is an inherently "optical" affair (Wollheim 1998). And it explains part of what is epistemically special about pictorial experience. Seeing a drawing or photograph of the Stonborough House can be a way of forming true, perceptually based beliefs about the building's visual appearance even in the absence of seeing the building itself. In short, even if pictures don't exhibit what we could call *subject transparency,* they often exhibit

property transparency.[19] And, since an object's visual appearance properties are presumably what we really care about when producing a picture of the object for representational, communicative purposes, property transparency does all of the epistemic work that philosophers like Lopes and Walton have credited to subject transparency.

A number of philosophers, including Richard Wollheim (1987, 1998, 2003), Dominic Lopes (1996, 2005, 2006), and Alva Noë (2012), have assumed that when we see and understand a picture of a thing belonging to a certain high-level kind *F*, for example, a painting of a plow, or a ship, or a shepherd, we experience an object with certain visible properties in pictorial space and, further, recognize that object as an *F*. Lopes, for instance, writes: "[Vermeer's] *Woman Holding a Balance* depicts a balance, and when you look at it, you typically have an experience as of a balance—you see a balance in the picture. [Degas's] *Woman with Field Glasses* depicts binoculars, and you typically see binoculars in it" (2005, 12). Similarly, when we see and understand a picture of an individual *a*, for example, a drawing of Alfred Hitchcock, or Secretariat, or the Great Pyramid of Giza, we experience an object with certain visible properties in pictorial space and, further, recognize that object as *a*. "[W]hen you look at a photograph of Hillary Clinton," Alva Noë writes, ". . . you see *her*. . . . Hillary confronts you when you see her picture. Hillary shows up for you, in your experience of the picture. She is present for you, visually, in the picture. Full stop. This is phenomenological bedrock" (Noë 2012, 83–84). We can call this assumption the (Pictorial) Recognition Thesis.

If pictorial experience isn't transparent to depicted objects, then—and this a second consequence of the deep resemblance theory—visually recognizing the object of pictorial experience cannot be a case of visually recognizing the depicted object itself. Recognition, unlike mere categorization or identification, is *factive*. You can mistakenly identify a woman in the distance or on a dark night as Hillary Clinton, but, if the woman isn't actually Clinton, then you cannot visually recognize her as Clinton, no matter what the viewing conditions are. In order to recognize Clinton, you need to *see* Clinton. In general, if the experience elicited by a picture is opaque to the individual it depicts, then the thesis that, when we look "into" a picture of an individual *a*, we see an object that we recognize as *a* is false. Full stop.

A similar point holds with respect to pictures of high-level kinds. When we look at a picture of an object belonging to a certain high-level kind *F*, according to the deep resemblance theory, we do not experience an object properly categorizable as an *F*. Rather, we experience a virtual object in pictorial space that models an *F*. Contrary to the Recognition Thesis, "That's a *horse*" asserted when viewing Stubbs's painting of Whistlejacket is like "That's a *horse*" asserted when viewing an equestrian bronze, or a topiary sculpture of a horse, or a toy horse. In all these

---

19. Compare Dretske's (1984) distinction between perceptual and informational transparency.

cases, it is the model's *original* that we are identifying as a horse rather than the model itself.[20]

Third, the deep resemblance theory reveals two important differences between pictorial and cartographic representational systems. The vehicles of cartographic representational content are composed of *coordinates* and *markers* (Rescorla 2009). Topological or distance relations between coordinates on the surface of a map reflexively represent topological or distance relations between locations in the world. Markers, by contrast, represent objects and properties, but typically do so by means of a conventional symbol of some kind. The location of a hospital, for instance, may be represented on a map by a red cross, a state capital by a star, population density by a certain arbitrary color, and a person by her name. Indeed, some types of maps make use of linguistic markers exclusively (Camp 2007, 158).

The first important difference between pictorial and cartographic representational systems is that pictorial content is entirely determined by reflexive representational elements. Pictorial representational content is *purely iconic*.[21] To be sure, not every experienced attribute of an iconic, virtual model need make contributions

---

20. Michael Newall has proposed that "Seeing something, *Y*, in a picture surface, *X*, involves the experience of seeing *X* and the non-veridical experience of seeing *Y*" (Newall 2009, 141). On this view, while categorizing the object we visually experience in a picture of a horse as a *horse* is necessary for pictorial understanding, this categorization is erroneous ("a false positive," as Currie 2009 puts it). The error, further, is one made by the visual system and not by the perceiver herself, who normally has access to "contextual cues" that specify the presence of the 2D pictorial surface (Newall 2009, 136). Call this the *nonveridical categorization theory*.

It is implausible that the visual system is typically "fooled" by pictures in the manner suggested by the nonveridical categorization theory. The properties attributed by the visual system to the object we experience in a horse-depicting painting, drawing, or etching, for example, typically differ from those attributed by the visual system to horses seen in the flesh, and they typically do so in *at least* as many ways as those attributed by the visual system to 3D sculptures and models of a horse. There is no reason to suppose, however, that the visual system is fooled when we see, say, an equestrian bronze or a toy horse under normal viewing conditions. And this is, in part, because our visual concept of the high-level kind *horse* is distinct from our visual concept of the high-level kind *sculpture of a horse*. Indeed, we are able to categorize a sculpture of a horse as such even when the sculpture has visible properties that no real horse could have and lacks visible properties that no real horse could lack.

The *dynamism* of our ability to recognize an object as belonging to a certain high-level kind *F* (Lopes 2005, 45–48; Newall 2009, 139), that is, our ability to identify an object as an *F* reliably across significant variation in the object's visual appearance, is irrelevant here since it only extends across variations in appearance that, from the standpoint of the visual system, are consistent with actually being an *F*. Just as the range of variations in visual appearance consistent with applying the concept *horse* must be distinguished from the range of variations in visual appearance consistent with applying the concept *sculpture of a horse*, they must also be distinguished from the range of variations in visual appearance consistent with applying the concept *picture of a horse*.

21. Which is not to deny that pictures can be used to make claims about the world that go well beyond their *pictorial* representational content. Pictorial representation, as Hopkins points out, is only one of several forms of representation that pictures exhibit (1998, 9). A picture, for example, may depict an object *a* that is conventionally symbolic of another object *b*, and by virtue of depicting *a* as having certain visual appearance properties, provide reason to believe some claim involving *b*.

to pictorial content. Which attributes do and which don't depends on the work that is done with the model, on how it is intended to be used by appropriate viewers or "consumers." The color of a virtual architectural model constructed using 3D, computer-aided design software, for example, need not stand for the color of the building that it represents. Those attributes of a virtual model that *do* contribute to pictorial content, however, contribute iconically. Cartographic representations, by contrast, are partly iconic and partly symbolic; they are hybrid representations that, in combining coordinates and markers, "balance direct resemblance and abstract conventionality" (Camp 2007, 159).

A second important difference is that cartographic representations, unlike pictorial representations, do not depend on the experience as of depth and 3D structure. The vehicles of pictorial representational content, unlike the vehicles of cartographic representational content, are virtual models constructed in pictorial space. Whereas maps represent depth, i.e., distance orthogonal to the ground plane, conventionally by means of numerals, contour lines, or colors, pictures represent depth iconically. When the Guidance and Systematicity requirements are met, depth and 3D structure in pictorial space can be used reflexively to represent depth and 3D structure in the world. Virtual models can be experienced as spatially isomorphic to their originals in three dimensions. By contrast, the vehicles of cartographic representational content, like the vehicles of written communication, are typically experienced as spatially flat or 2D.

A picture, for example, an aerial photograph of a city, can of course guide the way its viewers navigate from one place to another or form beliefs about the relative locations of certain landmarks. This observation doesn't blur the distinctions that I have drawn. It just means that a picture can sometimes perform some of the functions normally performed by a map. Further, while some maps, in particular, topographic maps that represent elevation by means of contour lines, can elicit experiences of virtual depth, this is merely incidental to how they perform their representational function. That one location on a mountainside is 300 meters higher than another is typically represented by the contour interval and the number of intervening contour lines (or simply by indices on the contour lines). That the map elicits an experience in which one location *looks* higher than another isn't necessary for the map cartographically to represent that relation as obtaining in the world.

Finally, although computationally early visual representations in the brain are often assumed to be topographically organized, image-like representations—Kosslyn et al. 2006, speak of "functional depiction"—there is an important disanalogy between conventional, pictorial representation and brain-based, visual representation. The vehicles of pictorial representation, on the present account, are virtual models that meet the Guidance and Systematicity requirements. They are the intentional objects of pictorial experience. There is no reason to suppose, however, that there is any analogue to pictorial experience in early visual processing. Higher-level visual, cognitive, and motor systems that receive early visual representations as

inputs are not homuncular perceivers. In general, we cannot explain visual representation in the brain in terms of depiction if depiction itself needs to be understood, fundamentally, in terms of space-representing, visual representation—in particular, visual representation of objects in phenomenally 3D, pictorial space.[22]

ACKNOWLEDGMENTS

REFERENCES

Abell, C. 2009. "Canny Resemblance." *Philosophical Review* 118: 183–223.
Ames, A. 1925. "The Illusion of Depth from Single Pictures." *JOSA* 10: 137–47.
Anderson, B., and J. Winawer. 2005. "Image Segmentation and Lightness Perception." *Nature* 434: 79–83.
Baker, G. 2001. "Wittgenstein: Concepts or Conceptions?" *Harvard Review of Philosophy* 9: 7–23.
Bantinaki, K. 2009. "Depiction." In *A Companion to Aesthetics*, 2nd edition, ed. S. Davies, K. Marie Higgins, R. Hopkins, R. Stecker, and D. Cooper. Oxford: Blackwell.
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. 2008. "Vision, Action, and Make-perceive." *Mind and Language* 23: 457–97.
———. forthcoming. "Gombrich and the Duck-Rabbit." In *Aspect Perception after Wittgenstein: Seeing-As and Novelty*, ed. M. Beaney and D. Shaw. London: Routledge.
Budd, M. 1992/2008. "On Looking at a Picture." In *Psychoanalysis, Mind and Art: Perspectives on Richard Wollheim*, ed. J. Hopkins and A. Savile, 259–80. Reprinted in Budd 2008.
———. 1993/2008. "How Pictures Look." In *Virtue and Taste: Essays on Politics, Ethics, and Aesthetics*, ed. D. Knowles and J. Skorupski, 154–75. Oxford: Basil Blackwell. Reprinted in Budd 2008.
———. 2008. *Aesthetic Essays*. Oxford: Oxford University Press.
Burge, T. 2005. "Disjunctivism and Perceptual Psychology." *Philosophical Topics* 33: 1–78.
———. 2010. *Origins of Objectivity*. Oxford: Oxford University Press.
Busey, T., N. Brady, and J. Cutting. 1990. "Compensation Is Unnecessary for the Perception of Faces in Slanted Pictures." *Perception & Psychophysics* 48: 1–11.
Camp, E. 2007. "Thinking with Maps." *Philosophical Perspectives* 21: 145–82.

---

22. For other objections to the view that early visual processing areas represent space by means of functional depiction, see Clark 2006.

Casati, R. 2010. "Hallucinatory Pictures." *Acta Analytica* 25: 365–68.

Clark, A. 2006. "How Do Feature Maps Represent?" Early Content Conference, University of Maryland, April 22, 2006.

Clark, A. 2013. "Whatever Next? Predictive Brains, Situated Agents, and the Future of Cognitive Science." *Behavioral and Brain Sciences* 36: 181–204.

Cohen, J., and A. Meskin. 2004. "On the Epistemic Value of Photographs." *The Journal of Aesthetics and Art Criticism* 62: 197–210.

Coulson, S., and G. Fauconnier. 1999. "Fake Guns and Stone Lions: Conceptual Blending and Privative Adjectives." In *Cognition and Function in Language*, ed. B. Fox, D. Jurafsky, and L. Michaelis. Palo Alto, CA: CSLI.

Cuijpers, R. H., A. M. Kappers, and J. J. Koenderink. 2000. "Investigation of Visual Space Using an Exocentric Pointing Task." *Perception & Psychophysics* 62: 1556–71.

Currie, G. 2009. "Art of the Paleolithic." In *A Companion to Aesthetics*, 2nd edition, ed. S. Davies, K. Higgins, R. Hopkins, R. Stecker, and D. Cooper, 1–10. Oxford: Wiley-Blackwell.

Cutting, J. 1997. "How the Eye Measures Reality and Virtual Reality." *Behavior Research Methods, Instruments, & Computers* 29: 27–36.

———. 2000. "Images, Imagination, and Movement: Pictorial Representations and Their Development in the Work of James Gibson." *Perception* 29: 635–48.

———. 2003. "Reconceiving Perceptual Space." In Hecht et al. 2003, 215–38.

Cutting, J., and P. Vishton. 1995. "Perceiving Layout and Knowing Distances." In *Perception of Space and Motion*, ed. W. Epstein and S. Rogers, 69–117. San Diego, CA: Academic Press.

Dretske, F. 1984. "Seeing through Pictures." *Noûs* 18: 73–74.

Feagin, S. 1998. "Presentation and Representation." *Journal of Aesthetics and Art Criticism* 56: 234–40.

Fleming, R., and B. Anderson. 2004. "The Perceptual Organization of Depth. In *The Visual Neurosciences*, ed. L. Chalupa and J. Werner. Cambridge, MA: MIT Press.

Franks, B. 1995. "Sense Generation: A 'Quasi-classical' Approach to Concepts and Concept Combination." *Cognitive Science* 19: 441–505.

Gentner, D., and A. Markman. 1997. "Structure Mapping in Analogy and Similarity." *American Psychologist* 52: 45–56.

Georgieva, S. S., J. T. Todd, R. Peeters, and G. A. Orban. 2008. "The Extraction of 3D Shape from Texture and Shading in the Human Brain." *Cerebral Cortex* 18: 2416–38.

Gibson, J. J. 1960. "Pictures, Perspective, and Perception." *Daedalus* 89: 216–27.

———. 1971. "The Information Available in Pictures." *Leonardo* 4: 27–35.

———. 1979. *The Ecological Approach to Visual Perception*. Boston: Houghton Mifflin.

———. 1980. "Foreword: A Prefatory Essay on the Perception of Surfaces versus the Perception of Markings on a Surface." In *The Perception of Pictures*, vol. 1, ed. M. Hagen, xi–xvii. New York: Academic Press.

Gilchrist, A. 1977. "Perceived Lightness Depends on Perceived Spatial Arrangement." *Science* 195: 185–87.

———. 1980. "When Does Perceived Lightness Depend on Spatial Arrangement?" *Perception and Psychophysics* 28: 527–38.

Godfrey-Smith, P. 2006. "Mental Representation, Naturalism, and Teleosemantics." In *Teleosemantics*, ed. G. Macdonald and D. Papineau, 42–68. Oxford: Oxford University Press.

Goldstein, E. 1987. "Spatial Layout, Orientation Relative to the Observer, and Perceived Projection in Pictures Viewed at an Angle." *Journal of Experimental Psychology: Human Perception and Performance* 13: 256–66.

Goldstone, R. 1994. "The Role of Similarity in Categorization: Providing a Groundwork." *Cognition* 52: 125–57.

Gombrich, E. H. 1951/1985. "Meditations on a Hobby Horse." Reprinted in *Meditations on a Hobby Horse and Other Essays on the Theory of Art*. Oxford: Phaidon. 1–11.

———. 1961/2000. *Art and Illusion*, 2nd rev. ed. Princeton, NJ: Princeton University Press.

———. 1972. "Illusion and Art." *Illusion in Nature and Art*, ed. R. L. Gregory and E. H. Gombrich, 193–243. New York: Charles Scribner's Sons.

Gonzalez, C., T. Ganel, R. Whitwell, B. Morrissey, and M. Goodale. 2008. "Practice Makes Perfect, but Only with the Right Hand." *Neuropsychologia* 46: 624–31.

Goodman, N. 1976. *Languages of Art*. 2nd ed. Indianapolis: Hackett.

Greenberg, G. 2013. "Beyond Resemblance." *Philosophical Review* 122: 215–87.

Grill-Spector, K., Z. Kourtzi, and N. Kanwisher. 2001. "The Lateral Occipital Complex and Its Role in Object Recognition." *Vision Research* 41: 1409–22.

Haugeland, J. 1991. "Representational Genera." In *Philosophy and Connectionist Theory*, ed. W. Ramsey, S. Stich, and D. Rumelhart, 61–89. Hillsdale, NJ: Lawrence Erlbaum Associates.

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

Hochberg, J. 1962/2007. "The Psychophysics of Pictorial Perception." *Audiovisual Communication Review* 10: 22–54. Reprinted in Peterson et al. 2008.

———. 1980/2007. "Pictorial Functions and Perceptual Structures. In *The Perception of Pictures*, vol. 2, ed. M. Hagen, 47–93. New York: Academic. Reprinted in Peterson et al. 2008.

Hohwy, J. 2013. *The Predictive Mind*. Oxford: Oxford University Press.

Hopkins, R. 1998. *Picture, Image, and Experience*. Cambridge: Cambridge University Press.

———. 2006. "The Speaking Image: Visual Communication and the Nature of Depiction." In *Contemporary Debates in Aesthetics and the Philosophy of Art*, ed. M. Kieran, 135–59. Oxford: Blackwell.

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.

Kamp, H. 1975. "Two Theories about Adjectives." In *Formal Semantics of Natural Language*, ed. E. Keenan. Cambridge: Cambridge University Press.

Kanwisher, N. 2004. "The Ventral Visual Object Pathway in Humans: Evidence from fMRI." In *The Visual Neurosciences*, ed. L. Chalupa and J. Werner. Cambridge, MA: MIT Press.

Kemp, M. 1989. *Leonardo on Painting: An Anthology of Writings by Leonardo da Vinci with a Selection of Documents Relating to His Career*. New Haven, CT: Yale University Press.

Kennedy, J. 1974. *A Psychology of Picture Perception*. San Francisco: Jossey-Bass Publishers.

Koenderink, J. J. 1998. "Pictorial Relief." *Philosophical Transactions of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences* 356: 1071–86.

———. 1999. "Virtual Psychophysics." *Perception* 28: 669–74.

———. 2012. *Pictorial Space*. Utrecht: De Clootcrans Press.

Koenderink, J. J., and A. J. van Doorn. 2003. "Pictorial Space." In Hecht et al. 2003, 239–300.

Koenderink, J. J., A. J. van Doorn, and A. M. Kappers. 1994. "On So-called Paradoxical Monocular Stereoscopy." *Perception* 23: 583–594.

Koenderink, J. J., A. J. van Doorn, A. M. Kappers, M. J. Doumen, and J. T. Todd. 2008. "Exocentric Pointing in Depth." *Vision Research* 48: 716–23.

Koenderink, J. J., A. J. van Doorn, and J. Wagemans. 2011. Depth. *i-Perception* 2: 541–64.

Koenderink, J., M. Wijntjes, and A. J. van Doorn. 2013. "Zograscopic Viewing." *i-Perception* 4: 192–206.

Kosslyn, S., W. Thompson, and G. Ganis. 2006. *The Case for Mental Imagery*. Oxford: Oxford University Press.

Kubovy, M. 1986. *The Psychology of Perspective and Renaissance Art*. Cambridge: Cambridge University Press.

Lopes, D. 1996. *Understanding Pictures*. Oxford: Clarendon Press.

———. 2003. "The Aesthetics of Photographic Transparency." *Mind* 112: 433–48.

———. 2005. *Sight and Sensibility: Evaluating Pictures*. Oxford: Oxford University Press.

———. 2006. "The Domain of Depiction." In *Contemporary Debates in Aesthetics and the Philosophy of Art*, ed. M. Kieran, 160–75. Oxford: Blackwell.

Machery, E. 2009. *Doing without Concepts*. 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.

Martin, M. G. F. 2010. "What's in a Look?" In *Perceiving the World*, ed. B. Nanay. Oxford: Oxford University Press.

———. 2012. "Sounds and Images." *British Journal of Aesthetics* 52: 331–51.

Mausfeld, R. 2003. "Conjoint Representations and the Mental Capacity for Multiple Simultaneous Perspectives." In Hecht et al. 2003, 17–60.

Medin, D. L., R. Goldstone, and D. Gentner. 1993. "Respects for Similarity." *Psychological Review* 100: 254–78.

Millar, B. 2006. "The Conflicted Character of Picture Perception." *Journal of Aesthetics and Art Criticism* 64: 471–77.

Millikan, R. 1989. "Biosemantics." *Journal of Philosophy* 86: 281–97.

———. 2000. *On Clear and Confused Ideas*. Cambridge: Cambridge University Press.

———. 2004. *The Varieties of Meaning*. Cambridge, MA: MIT Press.

———. 2012. "Natural Signs." In *How the World Computes: Turing Centenary Conference and 8th Conference on Computability in Europe*, ed. S. Cooper, A. Dawar, and B. Lowe, 496–506. Berlin: Springer Verlag.

Murphy, G. 2004. *The Big Book of Concepts*. Cambridge, MA: MIT Press.

Nanay, B. 2010. "Inflected and Uninflected Perception of Pictures." In *Philosophical Perspectives on Pictures*. ed. C. Abell and K. Bantinaki, 181–207. Oxford: Oxford University Press.

———. 2014. "*Trompe l'oeil* and the Dorsal/Ventral Account of Picture Perception." *Review of Philosophy and Psychology* 6: 181–97.

Neander, K. 1987. "Pictorial Representation: A Matter of Resemblance." *British Journal of Aesthetics* 27: 213–26.

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.

Newall, M. 2009. "Pictorial Experience and Seeing." *British Journal of Aesthetics* 49: 129–41.

———. 2015. "Is Seeing-in a Transparency Effect?" *British Journal of Aesthetics* 55: 131–56.

Noë, A. 2012. *Varieties of Presence*. Cambridge, MA: Harvard University Press.

Peterson, M. A., B. Gillam, and H. A. Sedgwick. 2007. *In the Mind's Eye: Julian Hochberg on the Perception of Pictures, Films, and the World*. New York: Oxford University Press.

Pirenne, M. H. 1970. *Optics, Painting, and Photography*. Cambridge: Cambridge University Press.

Ramscar, M., and U. Hahn., eds. 2001. *Similarity and Categorization*. Oxford: Oxford University Press.

Rescorla, M. 2009. "Predication and Cartographic Representation." *Synthese* 169: 175–200.

———. 2015. "Bayesian Perceptual Psychology." In *The Oxford Handbook of the Philosophy of Perception*, ed. M. Matthen. Oxford: Oxford University Press.

Rice, N. J., K. F. Valyear, M. A. Goodale, A. D. Milner, and J. C. Culham. 2007. "Orientation Sensitivity to Graspable Objects: An fMRI Adaptation Study." *Neuroimage* 36: T87–T93.

Rogers, S. 1995. "Perceiving Pictorial Space." In *Perception of Space and Motion*, ed. W. Epstein and S. Rogers, 119–64. London: Academic Press Limited.

———. 2003. "Truth and Meaning in Pictorial Space." In Hecht et al. 2003, 301–20.

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.

Schellenberg, S. 2013. "Experience and Evidence." *Mind* 122: 699–747.

Schier, F. 1986. *Deeper into Pictures*. Cambridge: Cambridge University Press.

Schlosberg, H. 1941. "Stereoscopic Depth from Single Pictures." *American Journal of Psychology* 54: 601–5.

Sedgwick, H. 2003. "Relating Direct and Indirect Perception of Spatial Layout." In Hecht et al. 2003, 61–76.

Shikata, E., F. Hamzei, V. Glauche, R. Knab, C. Dettmers, C. Weiller, and C. Büchel. 2001. "Surface Orientation Discrimination Activates Caudal and Anterior Intraparietal Sulcus in Humans: An Event-Related fMRI Study." *Journal of Neurophysiology* 85: 1309–14.

Taira, M., I. Nose, K. Inoue, and K. Tsutsui. 2001. "Cortical Areas Related to Attention to 3D Surface Structures Based on Shading: An fMRI Study." *Neuroimage* 14: 959–66.

Thompson, W., R. Fleming, S. Creem-Regehr, and J. K. Stefanucci. 2011. *Visual Perception from a Computer Graphics Perspective*. Boca Raton, FL: CRC Press.

Tsutsui, K., H. Sakata, T. Naganuma, and M. Taira. 2002. "Neural Correlates for Perception of 3D Surface Orientation from Texture Gradient." *Science* 298: 409–12.

Tsutsui, K., M. Taira, and H. Sakata. 2005. "Neural Mechanisms of Three-Dimensional Vision." *Neuroscience Research* 51: 221–29.

Vishwanath, D. 2011. "Visual Information in Surface and Depth Perception: Reconciling Pictures and Reality." In *Perception beyond Inference: The Information Content of Visual Processes*, ed. L. Albertazzi, G. van Tonder, and D. Vishwanath. Cambridge, MA: MIT Press.

———. 2014. "Toward a New Theory of Stereopsis." *Psychological Review* 121: 151–78.

Vishwanath, D., and P. Hibbard. 2013. "Seeing in 3-D with Just One Eye: Stereopsis without Binocular Vision." *Psychological Science* 24: 1673–85.

Wagemans, J., A. J. van Doorn, and J. J. Koenderink. 2011. "Measuring 3D Point Configurations in Pictorial Space." *i-Perception* 2: 77–111.

Walton, K. 1984. "Transparent Pictures: On the Nature of Photographic Realism." *Critical Inquiry* 11: 246–76.

———. 1990. *Mimesis as Make-Believe: On the Foundations of the Representational Arts*. Cambridge, MA: Harvard University Press.

———. 2008. *Marvelous Images: On Values and the Arts*. Oxford: Oxford University Press.

Weisberg, M. 2013. *Simulation and Similarity*. Oxford: Oxford University Press.

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.

Wiesing, L. 2010. *Artificial Presence: Philosophical Studies in Image Theory*. Stanford: Stanford University Press.

Wölfflin, H. 1929. *Principles of Art History*. Translated by M. D. Hottinger. New York: Dover Publications.

Wollheim, R. 1987. *Painting as an Art*. London: Thames & Hudson.

———. 1998. "On Pictorial Representation." *Journal of Aesthetics and Art Criticism* 56: 217–26.

———. 2003. "What Makes Representational Painting Truly Visual?" *Aristotelian Society Supplementary* 77: 131–47.