

CONTENT AND TARGET IN PICTORIAL REPRESENTATION

GABRIEL GREENBERG

University of California, Los Angeles

This essay argues for a model of pictorial representation which aims to explain the relationship between pictorial content and pictorial accuracy. Focusing on cases where pictures are intended to convey accurate information, the model distinguishes between two fundamental representational relations: on one hand, a picture *expresses a content*; on the other, it *aims at a target scene*. Such a picture is *accurate* when the content it expresses fits the target scene it aims at. In addition, the model follows the traditional division of content into two aspects: *singular content* specifies the particular individuals which a picture is *of*, and *attributive content* specifies the properties and relations which the picture ascribes to those individuals. For a picture to be accurate, both aspects must be matched in the target. I call this the *Three-Part Model* because it distinguishes between the triad of factors, singular content, attributive content, and target, which together determine pictorial accuracy. Though previous work on depiction has not recognized the distinctive role played by target, I will argue that it is essential in order to make sense of accuracy judgements across a range of central cases.

In Section 1, I introduce the the Three-Part Model. Section 2 refines the key definition of accuracy and defends the assumption that accuracy depends a contextually selected target scene. Then, in Section 3, I'll argue from cases that target scenes are independent of pictorial contents. Section 4 goes on to show how the Three-Part Model may be adapted to handle the phenomena of counterfactual and generic depiction. In Section 5, the conclusion, I suggest that the same three-part representational architectural extends to language, vision, and mental imagery.

Contact: Gabriel Greenberg <ggreenberg@humnet.ucla.edu>

1. Content and Target

Consider the following print, published in the early 1800's, as part of a project by French scholars to document what was known to them of ancient and contemporary Egypt.¹



Picture E

Picture E is first of all a picture of the Great Sphinx from the Giza Plateau on the bank of the Nile in Giza, Egypt. It is also a picture of the Pyramid of Khafre (on the left) and the Pyramid of Khufu (just visible on the right). We may also suppose that the horse and rider pictured in front of the Sphinx were drawn from life; it is a picture of them as well. Not only is Picture E *of* various objects, it also depicts them *as* having a variety of features. Thus it depicts the Sphinx as having a certain shape (e.g. with no nose), as sitting in a certain position relative to the pyramids, as catching the light at a certain angle. It depicts the Pyramid of Khafre as having a certain shape, as sitting in a certain position relative to the Sphinx. And so on.

What a picture is *of*, and what it depicts its subjects *as*— these are reflections of a picture's **CONTENT**. Content corresponds roughly to *what's happening* in the picture, or how the picture *depicts* the world, independent of whether the world fits this construal. And pictorial content determines substantive accuracy conditions. When these conditions are met, the picture is **ACCURATE**, and when

1. From: *Description of Egypt or Collection of Observations and Researches Made in Egypt during the French Army Expedition*. Paris: Imperial Printing, 1809-1813, Royal Printing, 1817- [1830]. Volume V, plate 12.

they are not, it is *INACCURATE*, it misrepresents.

Over the last fifty years, following Goodman (1968), scholars of depiction have distinguished between two aspects of pictorial content. These have gone by many names, but for the sake of standardization, I'll call them *SINGULAR CONTENT* and *ATTRIBUTIVE CONTENT*.² Singular content includes all the individuals a picture is *of*— in the case of Picture E, the Sphinx, the pyramids, the horse, the rider, and so on. Attributive content includes all the properties and relations ascribed to those individuals— here, their shape, orientation, illumination, and so on, and possibly also high-level properties like *being a statue* or *being a person*. Together, singular and attributive content make up an integrated whole.³ To a first approximation, when a picture depicts some object *X* as having some property *F*, then the object in the *X* position is part of the singular content, while the property in the *F* position is part of the attributive content.⁴

In determining a picture's content, the picture's own spatial and chromatic organization plays the primary role; no content, attributive or singular, is expressed by a picture except through some particular region and pattern of marks on the picture plane. Still, the picture itself only gains semantic significance within a

2. Goodman (1968: 27-8) distinguishes the *kind* of a picture from its *denotation*. Kaplan (1968: 197-8) differentiates between a picture's *descriptive content* and its *genetic character*. Kjörup (1974: 220) marks a split between *predication* and *reference* in depiction. Schier (1986) distinguishes *iconic predication* from *iconic reference*. Hyman (2012: 136) separates out a picture's *sense* and its *reference*. And in Greenberg (2013: 222), I mark a distinction between a picture's *content* and its *referent*. Note that not all authors have treated (what I am calling) *singular content* as part of (what they call) "content." Instead, "content" (or cognates, like "sense") are sometimes reserved for what I here call *attributive* content, while an independent semantic relation ("denotation", "reference") is posited for the expression of what I here call *singular* content (e.g. Greenberg 2013, Hyman 2012).

3. Why think that singular and attributive aspects of pictorial representation together constitute a unified pictorial content? First, facts about what objects pictures are *of* and facts about what those objects are represented *as* seem to be thoroughly intermingled. It's not as if a picture is *of* Cubey, and then, separately, depicts something *as* a cube. Instead, it depicts *Cubey* as a cube. Second, judgements of accuracy support the conclusion that there is at least some level of pictorial content in which what a picture is of determines accuracy conditions. Suppose I draw an accurate picture of Cubey sitting on my desk. My other favorite cube, Yebuc, might be sitting on another desk in the same configuration. An accurate picture of Yebuc might come out qualitatively identical to my picture of Cubey. Still there is a sense in which what is depicted by the original picture— that picture's content— would *not* hold at the scene in which Yebuc is visible. This is simply because the content contains the object Cubey itself. (I don't deny that we may also, more or less at will, track purely attributive content as well. The point is that there is a level of content which incorporates singular content.)

4. For expository purposes, I set aside non-referential aspects of singular content throughout this essay. Thus I don't discuss modes of presentation, object senses, or other singular hyper-intensional contents. Peacocke (1992) and Burge (1991; 2014), among others, have argued for such elements in visual perception, and they likely arise in depiction as well. In addition, I set aside cases of *indefinite* content, as when a picture merely depicts *some* cube as being located in a given direction, rather than a particular cube.

CONTEXT, understood as the particular historical, causal, social, and psychological setting in which it is created.

Here I assume a broadly contextualist view of content determination. Let a PICTURE be a two-dimensional image token. I'll say that a picture is CREATED in a context, and thereby EXPRESSES its content relative to the context in which it was created.⁵ In this capacity, context plays two roles. On one hand, context determines an operative SYSTEM OF DEPICTION, the pictorial analogue of a language. Systems of depiction play an essential role in associating the geometrical and chromatic surface features of pictures with elements of attributive content (Giardino & Greenberg 2015).⁶ On the other, following Kaplan (1968: 197-8) and Lopes (1996: ch. 5), I assume that a picture's singular content is largely a function of artist intentions, meditated by the causal context in which the picture was created. In this way, pictorial singular content seems to be fixed in a manner at least analogous to that envisioned by the causal theory of reference for names.⁷

Much of the literature on depiction can be understood as offering accounts of the way in which pictures express their contents. For example, recent work on the resemblance theory (Neander 1987; Abell 2009; Blumson 2014), structural approaches to depiction (Hyman 2006; Kulvicki 2006; Greenberg 2013) and per-

5. In the case of drawing or painting, the context of creation is the same context in which the token picture is physically realized. This is not the case for photography or printmaking, where the context of creation (and the context relevant for determining content) may be temporally and causally antecedent to the context in which the picture is physically realized. In the latter cases, we may say (a bit oddly) that a picture was *created* in context *c*, but did not express content in *c* until it was *physically realized* at another, later context. I gloss over this complication in what follows.

6. Ultimately it may be necessary to relativize content expression to a broader suite of SYSTEMS OF INTERPRETATION, which capture the general rules by which picture types are mapped to all aspects of their attributive content. Such systems are likely heterogeneous, including both visual and pragmatic mechanisms, as well as purely geometric and chromatic constraints (Kulvicki 2006; Abell 2009). It is the last type of system which I term *systems of depiction* here.

7. Beyond these rough generalizations, different authors apportion different aspects of content to systematic rules, intentional and causal connections, and contextualized pragmatic inference respectively. One approach, suggested by Goodman (1968: ch. 1), Kjørup (1978: 57), and Hyman (2012: 138-40), holds that that all attributive content is determined by general and systematic interpretive rules, while singular content traces more directly back to local intentional and causal features of the context. A different perspective, defended by Kulvicki (2006), is that only a core "bare-bones" aspect of spatial attributive content is fixed by systematic rules, while the remaining "fleshed out" attributive and singular content are worked out pragmatically in context. A more extreme position is articulated by Abell (2009) who holds that nearly all aspects of content are determined by intention and general pragmatic capacities, without any recourse to specific rules of interpretation. In what follows, the judgements I elicit reflect a commitment to both systematic rules and for intention and causal connections, in the style of Hyman (2012). But officially, the Three Part Model requires only that pictures express their singular and attributive content relative to a context, imposing no specific constraints on the division of interpretive labor.

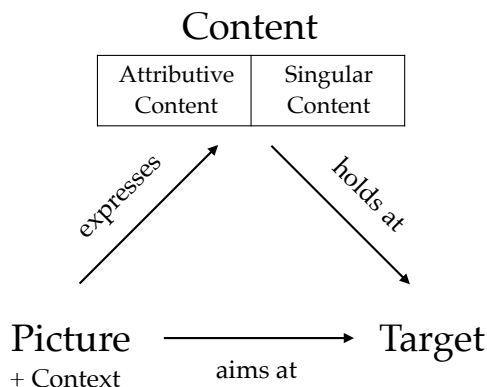
ceptual or experiential theories (Peacocke 1987; Lopes 1996; Hopkins 1998; Newall 2011) all take aim at the same core problem: what constraints guide the mapping from picture to the content it expresses? In this essay I sidestep this important debate by simply assuming that pictures express their contents in one way or another.

Allowing that pictures do express content, my agenda here is to ask about a distinct but central feature of pictorial representation: what makes a picture *accurate* or *inaccurate*? While this question has received little direct attention, a certain type of answer is discernible in the background of much contemporary philosophy of depiction, and it is this account that I will contest below. Very roughly, this is the idea that whether a picture is an accurate depiction or not is determined entirely by the degree of fit between two factors, a picture's singular content and its attributive content (measured against the backdrop of the state of the actual world, in some accounts). It is in this spirit that Goodman (1968: 38) writes, "for a picture to be faithful [\approx accurate] is simply for the object represented to have the properties that the picture in effect ascribes to it."⁸ The same assumption is perpetuated in recent formal theories of depiction, like my own account of pictorial accuracy in Greenberg (2013: 252), or Casati and Varzi (1999: 194-5) and Rescorla (2009: 180) on maps— where accuracy (or truth) is explicitly defined in terms of singular and attributive factors alone. In what follows I'll argue that this two-part approach reflects an impoverished conception of the relationship between pictorial content and accuracy, and as a consequence cannot account for a wide range of critical accuracy judgements.⁹

The novelty of the Three-Part Model is the counter-claim that accuracy is a function only in part of singular and attributive content— but also of a further, contextually selected index that I will call, following Cummins (1996), the *TARGET* of the picture. The interplay of picture, content, and target envisioned by the Three-Part Model is illustrated schematically below:

8. Or consider Kjørup (1978: 62-64), characterizing the artist's act of accurate depiction: "the producer of the picture must apply the picture to some referent and predicate something about the referent through the picture."

9. The Three-Part Model advocated below should not be confused with Nanay's (2018) "three-fold" approach to picture perception (nor the two-part approach with Wollheim's (1987) "two-fold" account of picture perception). The Three-Part Model is a theory of the determiners of pictorial accuracy. Nanay's three-fold approach, by contrast, is a theory of picture perception.



The Three-Part Model

In brief: Let P be a *picture* created in c . Then, in c , P *expresses* a content, with both singular and attributive aspects. And in c , P *aims at* a target scene. Finally, the content of P may or may not *hold at* the target of P . The key definition of accuracy is formulated in terms content and target: in c , P is *accurate* if and only if the content that P expresses in c holds at the target that P aims at in c . Each of these elements are explained below, with a more explicit definition of accuracy taken up in the next section.

In the Three-Part Model, the target of a picture is a contextually selected index of evaluation. It takes the form of a possible spatio-temporal situation anchored at a particular viewpoint, what I will call a *SCENE*. I'll model scenes as *VIEWPOINT-CENTERED WORLDS*, that is, as pairs of worlds and viewpoints. A *VIEWPOINT* here is conceived of as an oriented location, situated in space and time at a particular world, and carries no implication of a real or metaphorical viewer. I'll think of pictorial content as *HOLDING AT* a scene, much the way that propositions are thought to hold at worlds (Ross 1997; Blumson 2009; Abusch 2015). The fact that scenes include an index for viewpoint reflects the fact that a given pictorial content might hold at the actual world relative to *one* viewpoint (or location), but not at another. (See Section 2.) Crucially, since content has both singular and attributive aspects, for a given content to hold at a scene, both singular and attributive aspects of the content must be matched in the scene.

According to the Three-Part Model, for a picture to be accurate, its singular content must instantiate its attributive content, *in the target scene*. When an artist sets out to create a drawing, she comes to the table with at least two kinds of intentions. One is the intention to create a picture which expresses a particular content. The other is an intention to create a picture whose content, whatever it may be, is accurate at a particular scene. In effect, the *target* is the subject matter of a picture, the *content* of the picture provides a kind of comment, and it is the intended function of the picture to offer that content as a comment on the target.

Consider again the case of Picture E. The content of E is apparent—it describes a certain space, populated, as I have suggested, with objects like the Sphinx, the pyramids, and the rider, which are in turn attributed properties of shape, distance, texture, illumination, and so on. Because E was designed to document a particular scene, we know that its target consists of a particular time (in the 1820's), a particular location (on the Giza Plateau in Egypt), and a particular oriented viewpoint within that location. Picture E is accurate to the extent to which its target happens to realize the singular and attributive content which it expresses.

Here it should be noted that the Three-Part Model is not intended to apply to *all* forms of pictorial representation. Pictures are used in a variety of ways. A central class of uses employ pictures to convey accurate information about or to “depict” the world; these include scientific or factual illustrations like Picture E above, newspaper photographs, life drawings, and much more. I shall call these *ASSERTORIC* uses of pictures (Kjørup 1978; Eaton 1980; Korsmeyer 1985). It is natural to evaluate assertoric pictures for accuracy, and it is such pictures which the Three-Part Model associates with targets. In this sense, the model can be thought of as one component of a more general *theory of pictorial acts* (Novitz 1975; Kjørup 1974; 1978).

By contrast, *IMPERATIVAL* uses of pictures, like Ikea instructional diagrams or road-side warning signs, function to convey instructions or plans, but not to be accurate per se. Still other kinds of images, like doodles, patterns, and some artworks, are neither assertoric nor imperatival; their central function is to please, inspire, stir the imagination, or trick the perceptual system. In all these cases, it seems unnatural, if not a kind of conceptual mistake, to ask whether such pictures are accurate.¹⁰ Since the target of a picture is the scene relative to which it functions to be accurate, pictures which do not aim at accuracy in this sense do not have targets.¹¹

The Three-Part Model, then, is directed only to assertoric uses of pictorial representation, pictures which may be appropriately evaluated for accuracy. Just as assertion plays a central role in the study of language, understanding assertoric depiction is central to the study of depiction. Henceforth, when I refer to “pictures” or “pictorial representation” without further specification, I mean to restrict my attention to cases of assertoric depiction.

The target of a picture is that scene which it is the picture's function to be

10. Though it is possible, in a derivative sense, to ask by fiat of an imperatival picture, or any other type of picture, whether it *would* be accurate relative to an arbitrary situation.

11. Adjudicating this question in particular cases is delicate. While some pictures undoubtedly lack targets, others (including many artworks and childrens' drawings) function to *introduce* or *specify* an imagined scenario. Such pictures are, arguably, *trivially* accurate, for they *have* targets—the very scenes they served to introduce or specify. In general, it is more natural to assess pictorial contents for accuracy relative to scenes which have already been specified in prior discourse or mental activity.

accurate at. (Officially, the target is the scene which the picture functions to be accurate at *or* to be inaccurate at, as in the case of a pictorial lie.) Thus a picture which is accurate of its *target* has achieved an important standard of representational success not conferred by mere accuracy at *some* scene. If I set out to draw a picture of my office from the viewpoint of the doorway, thereby fixing my target, and if the picture is accurate at this scene, it successfully fulfills its representational function; by contrast, if it is inaccurate at this scene, but happens to be accurate of some other scene, say in some other office, the picture has not succeeded as a representation.

The notion of representational *function* here is broad.¹² For many kinds of pictures, it is fixed by the artist's intentions or purposes: thus life drawings are intended to accurately represent the scene immediately before the eyes of the artist, so this is the scene picked out as the target. In the case of mechanically produced images, like digital photographs, it may be the function of the picture-taking device, rather than the intentions of the artist, which fix the target. Typically, in these cases, the target is the scene before the lens of a camera. In other cases, the causal relation between picture and target is less direct. In drawing from memory, the artist intends her picture to be accurate at some past scene which was *previously* before her eyes; then that past scene is the picture's target. In still other cases, the fixation of target need not be mediated by the artist herself seeing the target scene at all: in a police sketch, for example, it is the witness who sees the original scene and only verbally reports it to the artist; but since the picture is intended to be accurate at the originally witnessed scene, that is the target.

It is not even necessary that the target of a picture be seen by *anyone*. Thus, I may set out to draw the Sphinx from a bird's-eye view, based on my background knowledge of the terrain; though I have never occupied that viewpoint (or talked to someone who did), it still defines the target scene for my picture. Even more extreme, targets may be located in the future, as when I set out to draw (what I expect will be) the state of the Sphinx in 100 years; then the target is located at that future time. Or, if you ask me what it *would* look like if an asteroid were to collide with the Pyramid of Khafre, and I draw a picture in response, then the target of my picture is a counterfactual scene— one that has been only imagined, but cannot be viewed. As these cases show, just as representational intentions can range across time, space, and possibility, so too can targets. Perhaps the further a target is from the artist's immediate visual context, the more likely the picture is to be inaccurate. But no matter for target, whose role is simply to fix the standard by which accuracy is measured.

Central to the Three-Part Model is the conceptual distinction between *content* and *target*. In context, content and target are also determined by very different

12. The concepts of representational function and success here derive from Burge's (2010: 308-315) discussion of representational function in perception.

kinds of factors. A picture's content is always grounded in the spatial and chromatic organization of the picture itself; while intentions and other contextual factors play a role in determining content, they only do so via features of the picture plane. By contrast, the target of a picture is unconstrained by the format of the picture itself, and may be entirely fixed even before the artist begins work. Though both content and target are determined in part by what I am calling *context*, they differ with respect to the features of context they are determined by. As a consequence, there is no guarantee that a picture's content will fit its target, no matter how earnestly an artist intends to depict accurately. When such intentions are not realized, and content and target come apart, inaccuracy results.¹³

The independence of the determiners of content and target mean that content and target themselves are independent. Two pictures may have different contents, but the same target: perhaps *two* artists attempted to capture the scene in front of the Sphinx that day in the 1820's. One produced Picture E; the other artist, much less skilled, produced something wildly inaccurate, call it Picture E'. Clearly, Picture E and E' have different spatial contents; but in another sense, they seem to depict the same thing; this is the sense in which they have the same target—they aim at the same location, time, and viewpoint in Giza, in front of the Sphinx. One is accurate and the other inaccurate precisely because they have different contents but the same target.

Here it is important to distinguish the target, conceived as the intended index of evaluation for a picture, from *the content the artist intended to express* with the picture. Both are, in some sense, ideals of picture production— one is an ideal of expression, the other of evaluation— but the two are independent. Consider an artist with prodigious artistic skill, but whose memory is unreliable, and who sets out to draw a particular scene from memory. Because of her artistic skill, the content her picture expresses is exactly what she intended to express by it; but because of her faulty memory, that content may be highly inaccurate at the scene she intended to draw. Thus the target of a picture and the intended content of a picture come apart in characteristic ways. Given an artist's expressive intention, whether a picture expresses the content it was intended to is largely a matter of artistic ability, and is independent of how the world is. But given a picture's content, whether that content is accurate at its target is wholly a matter of how the world is (at the target), and is independent of artistic skill.

Though content and target are distinct, they can be confused. This is due in

13. Here I assume that both content and target are fixed by the context of creation. In unusual cases, a picture may be *repurposed* to aim at a new target. For example, a life drawing of a particular eagle might find its way into an encyclopedia, where the picture is used to depict eagles in general. (See Section 4 for more on generic depiction.) In such cases, a more complex notion of context would be required.

part to an ambiguity in talk of “depiction” and its cognates. Goodman (1968: 22) has already observed that speaking about what a picture “depicts” or what it is “of” is ambiguous between ascriptions of attributive and singular content. To add to the interpretive possibilities, one may use the same expressions to refer to a picture’s target as well. It is in this sense that Picture E is *of* a particular view of the Giza Plateau at a certain moment in the 1820’s. Both content and target correspond loosely to “what a picture depicts”.

For the most part, the philosophical literature which seeks to give necessary and sufficient conditions on “depiction” is directed at one or another aspect of pictorial content, and has not dealt directly the concept of target.¹⁴

More direct antecedents can be found within the philosophy of language and philosophy of mind. The terminology of “content” and “target” itself is adapted from Cummins’ (1996) theory of mental representation, though the correspondence between his view and the present account is only approximate.¹⁵ Closer relatives emerge from contemporary philosophy of language. There, any number of authors distinguish the *proposition* expressed by a sentence (\approx content)—containing both predicative and referential constituents—from an element relative to which the proposition is true or false (\approx target); this element is variously characterized as the *index* (Lewis 1980), *world* (Kripke 1972), or *circumstance of evaluation* (Kaplan 1989) for a sentence, in context. The alternative tradition of *situation semantics* offers an even closer parallel to the Three-Part Model. Following Austin (1950), sentences are evaluated for truth relative to *topic situations*, contextually selected parts of possible worlds, which in many ways resemble targets (Barwise & Etchemendy 1987; Kratzer 2017). Indeed, Kratzer (2017) argues that the dual structure of content and topic may apply to a wide variety of representations and propositional attitudes beyond those of Austin’s original concern. I explore such generalization in the conclusion.

14. An exception is Kjørup (1978) who takes “depiction” to be a special pictorial act which aims at accuracy, roughly equivalent to what I call “assertoric depiction.” My own method is to use the term “depiction” loosely, identifying and debating more specific representational relations as required, with the caveat that I normally use “depiction *of*” in the singular content sense.

15. Cummins’ notion of a target seems to be that of the *content* which a computational system is supposed to express in context; by contrast, my notion is that of the *index* relative to which content is supposed to be evaluated for accuracy. In addition, Recanati (2001) uses the term “target” in his account of quotations as verbal depictions. But his use of “target” more nearly means the pictorial referent, or singular content, of the quotation. It is taken to be a token utterance, i.e. an individual, and it is contrasted with the attributive content of the quotation.

2. Target and Accuracy

What distinguishes the Three-Part Model from previous work on depiction is its target-theoretic definition of accuracy. In this section, I'll first make this definition explicit, and then argue that we should accept it over an alternative definition of accuracy which dispenses with the notion of target altogether.

To fix ideas, I'll begin with a notion of relative accuracy: a picture, in context, is ACCURATE AT a scene when the content it expresses, in context, holds at that scene. (Pictures are only accurate at a scene *in a context* because pictures only express content in context.) A pictorial content holds at a scene when both the singular and attributive components of the content correspond in the appropriate way with the objects, properties, and relations which populate that scene. Thus, in the Three-Part Model, for a picture in context to be accurate at a scene requires (i) that the objects in the singular content actually exist in the scene; and (ii) that the objects so depicted have the properties and stand in the relations, in the scene, that they are associated with by the picture's attributive content.

An important wrinkle here is that accuracy, unlike truth, comes in degrees; what degree of accuracy defines a picture's accuracy conditions? Here I help myself to the notion of maximal or PERFECT accuracy. Henceforth, by "accuracy" I mean perfect accuracy (or very near it); by "inaccuracy" I mean less than perfect accuracy. Providing an account of graded accuracy is left to future investigation.¹⁶

Accuracy, in the sense I intend, does not imply realism or closeness to reality. A black and white drawing and a color painting may each be perfectly accurate, albeit in different systems, though the experiences these pictures elicit obviously differ in their proximity to normal perceptual experience. A full scale, working model might be closer to reality, in some sense, than a technical drawing, but each may be perfectly accurate. What matters for accuracy is the absence of misrepresentation— not the quantity or type of information represented.¹⁷ A picture in context is accurate at a scene when its content holds at that scene. That is, it is accurate at a scene to the extent that when it represents things as being a certain way in that scene, that is the way those things are.

16. In this characterization we risk losing sight of that aspect of pictorial content which, relative to a scene, determines intermediate degrees of accuracy. Still, it should be noted that those scenes at which a picture is perfectly accurate are, intuitively, exactly those which reflect its content. For example, given a drawing of my plant, it is part of the content of the drawing that it has precisely *that* shape of leaf— not big spiky leaves, not even leaves which are slightly more spiky, nor leaves which are slightly less spiky. For any leaf-shape not perfectly accurately represented by the picture, that shape is not part of the content of the picture. For these reasons, perfect accuracy rather than graded accuracy seems to play the foundational role in an account of pictorial content.

17. This does not mean that pictures may be automatically accurate in virtue of being blank or omitting marks— depending on the operative system of depiction, blankness itself can carry content. (See Camp 2007 and Rescorla 2009 for discussion.)

So far I've discussed the ways in which a picture in context may be evaluated for accuracy relative to an arbitrary scene. The concept of *target* serves to isolate one such scene as having special status: a picture's target corresponds to the scene relative to which the picture *functions* to be accurate. My proposal is that a picture, in context, is accurate *simpliciter* when it is accurate *at* its target.

Note that context plays two roles here, just as it does in a traditional model of linguistic meaning (Kaplan 1989). On one hand, context determines the factors necessary for a picture to express its content. (This is what MacFarlane (2009) calls *indexical context-sensitivity*.) On the other hand, context fixes the target which the picture aims at, thereby securing the scene relative to which the picture is accurate or inaccurate; such context-sensitivity effects the accuracy value of a picture, but not its content. (This is what MacFarlane calls *non-indexical context-sensitivity*.)

In context then, a picture both expresses a content and aims at a target. Given a context, when a picture is accurate *at* its target then it is accurate *simpliciter*. Combining this formulation with the characterization of relative accuracy offered above, we may derive the following statement of accuracy conditions for pictures. For any assertoric picture *P* and context *c*:

Accuracy: Three-Part Model

P is accurate in *c* if and only if

the attributive content expressed by *P* in *c*
is instantiated by the singular content expressed by *P* in *c*
in the target scene selected by *c*.

This definition is the hallmark of the Three-Part Model. In this section and the next, I make the case for this definition in stages. First, in the remainder of this section, I argue that whether a picture is accurate depends in part on a contextually selected scene. Then, in the next section, I'll argue that such a scene cannot be part of, or derived from, a picture's content. For reference, I set out both theses here, restricting to cases of assertoric depiction:

Thesis 1. A picture's accuracy depends in part on a contextually selected target scene.

Thesis 2. The target scene for a picture is independent of its content.

Together, Theses 1 and 2 support the Three-Part Model's definition of accuracy. Thesis 1 establishes the relevance of a target scene to determinations of accuracy, but falls short of claiming that this index is not a part of, or derived from, a picture's content (or vice versa). Thesis 2 goes on to separate the target scene, so construed, from content. Putting these together yields a definition of accuracy which necessarily adverts to both content and target as separate parameters, as in the formulation above.

I turn now to defend Thesis 1. This thesis distinguishes the Three-Part Model from a prominent interpretation of the two-factor approach to pictorial accuracy to be found in the traditional literature on depiction. According to that view, a picture is accurate if and only if its singular content instantiates the properties and relations in its attributive content *in the actual world*. Such theories pursue what I will call the ACTUALITY APPROACH, articulated by the accuracy-conditions below.¹⁸

Accuracy: Actuality Approach

P is accurate in *c* if and only if
 the attributive content expressed by *P*
 is instantiated by the singular content expressed by *P*
 in the actual world (of *c*).

The actuality approach contrasts with Thesis 1. For one, the actuality approach fixedly ties accuracy to the state of the actual world, whereas according to Thesis 1, the world of the target must be allowed to vary by artist intention— this is the sense in which the thesis requires that the target be *contextually selected*. Second, the actuality approach makes no room for variation of viewpoint in the index of evaluation, but Thesis 1 makes this an essential feature of the target, in virtue of characterizing it as a *scene* rather than a world.

The first step in arguing for Thesis 1, then, is simply establishing that assessments of pictorial accuracy depend on an implicitly selected world and temporally-located viewpoint; from there I'll make the case that this index is selected by context in the specific manner of targets.

A basic motivation for such scene dependence is the observation that the content of a given picture may hold or fail to hold at arbitrary scenes. Thus I may ask of Picture E whether what it depicts— its content— *would* hold at various alternative points in time and space. Intuitively, there is variation here. Variation with respect to *world*: as things actually are, the content of the picture holds; but in a counterfactual situation in which the Pyramid of Khafre had never been built, the same content would not hold. Variation with respect to *time*: at some points in time the Sphinx lacked a nose, and the content of Picture E holds at these times; at other times, it had a nose, and the content does not hold at those. And variation with respect to *viewpoint*: though the content of Picture E may hold at some viewpoints, at others, where the Pyramid of Khafre is not visible for instance, that content would not hold. Thus pictorial contents hold at conjunctions of worlds, times, and viewpoints— in other words, at scenes.

18. In the final clause of the definition below, I employ the ambiguous phrase “the actual world (of *c*)”; the most general way to understand this clause, following Kaplan (1989), is as *the world of the context c*, rather than tying the condition rigidly to the actual world. This means that we may meaningfully ask after the accuracy of pictures created in hypothetical situations.

The same lesson emerges from considering pictures produced in different contexts. Recall the contrast between actual and hypothetical depiction. On one hand, if you ask me to draw a picture of Saturn in orbit, and I produce a picture in response, whether my picture is accurate would depend on the state of the actual world. On the other, if you ask me to draw what it *would* look like if Saturn and Jupiter collided, whether my picture is accurate depends on the state of a counterfactually specified situation, and not directly on the actual world. The same type of variation emerges for viewpoint: if you are asked to draw a picture of what is in front of you, the accuracy of the picture depends on your current viewpoint; if you attempt to draw a familiar landmark from a bird's eye view, the accuracy of the picture now depends on the imagined viewpoint suspended in air. And parallel reasoning applies to variation in times. Thus different contexts have the effect of making pictorial accuracy depend on different worlds and viewpoints, just as Thesis 1 predicts.

The conclusion that a picture can only be assigned an accuracy value relative to an implicitly specified viewpoint-centered world is all but inevitable from the perspective of possible world semantics. For, according to a standard possible-worlds framework, what properties an object instantiates depends on the world under consideration. Similar dependencies emerge in relation to times and viewpoints. What properties an object instantiates varies by time; the Sphinx had a nose at one time, and not at another. And the distinctively perspectival properties and relations attributed by a picture are only instantiated relative to a viewpoint. Thus the Sphinx occludes the Pyramid of Khafre (and not Khufu) relative to some viewpoints; the reverse is true relative to others. Putting these elements together— world, time, and viewpoint— one arrives once again at the concept of a target scene.

Such considerations support the claim that pictorial accuracy depends on an implicitly specified scene. But which scene in particular is relevant for determining whether a picture, in context, is accurate *simpliciter*? The view of the Three-Part Model is that this scene is selected by context, in a manner largely dependent on the artist's intentions (or the function of the camera).

In certain respects, this position follows a mainstream view in philosophy of language, descendent from Kaplan (1989). In such a semantics, sentences are tokened in a context; relative to a context they express a proposition; and relative to a context they are assigned a circumstance of evaluation. This much is held in common with the Three-Part Model. The two views differ about the *way* in which context determines the circumstance of evaluation. Kaplan assumes that, for unembedded sentences, the circumstance of evaluation is determined by a simple default rule which privileges the world of the context over others: ϕ is true in c if and only if the proposition expressed by ϕ is true at the world of c . By contrast, in Three-Part Model, the target is flexibly determined by artist intentions.

This conclusion, I believe, is forced on us by cases.

The simple rule equating circumstance of evaluation with the actual world (or world of the context) cannot hold for the pictorial case, for two reasons. First, the relevant world of evaluation is not always the actual world. We saw this already for cases of hypothetical depiction: in some cases accuracy doesn't depend on the state of the *actual* world, but rather the state of some *counterfactual* scenario. In general, as the cases of future, past, and hypothetical depiction illustrate, the world and time of evaluation for a picture cannot be identified with the world and time of the context of creation. Instead, they correlate more directly with the representational intentions of the artist.

Second, there is the matter of viewpoint. There is no metaphysically privileged viewpoint relative to which pictures should be evaluated in the same way that the actual world might be considered a privileged world of evaluation. Viewpoint must either be selected for by context, or quantified over in the definition of accuracy. Consideration of cases suggests that it must be the former. Suppose I set out to draw the Sphinx from an angle where it in fact occludes the Pyramid of Khafre (as in Picture E). But what results, Picture E*, is a picture in which the Sphinx does *not* occlude the Pyramid of Khafre. We may also suppose that Picture E* is accurate from some *other* viewpoint besides the one intended. While there may be a secondary sense in which Picture E* correctly depicts the target scene, it is not in the first place a successful representation, for it fails to meet the representational standard set for it by the artist. It is not accurate because its content does not hold at the intended viewpoint. Generalizing, it seems that, like the world of evaluation, the viewpoint of evaluation is also picked out in context largely through the intentions of the artist.

I've argued that pictorial accuracy depends on an implicitly specified scene; and I've argued, in addition, that this scene must be selected in context in large part by the intentions of the artist (or the function of the camera). This conclusion establishes Thesis 1: a picture's accuracy depends in part on a contextually selected target scene.

3. Target and Singular Content

In the last section I argued that pictorial accuracy depends on a target scene. This in itself leaves open the question of what connection, if any, holds between a picture's target scene and its content. In this section, I argue for Thesis 2, the claim that target is independent of singular content and attributive content. Just as Thesis 1 contrasted with the actuality approach, Thesis 2 contrasts with a different set of strategies for reviving the traditional, two-factor conception of accurate depiction. These views accept that a contextually selected scene plays an essential role in determining accuracy, but attempt either to assimilate the objects

a picture is *of* into target (and do away with singular content), or to assimilate the scene into singular content (and do away with target). Either way, these theories countenance only *two* representational relata as the determiners of a pictorial accuracy, as against the central posit of the Three-Part Model.

Such two-factor views are spurred by the fact that singular content and target are easily confused. For, according to the Three-Part Model, pictures now have two relata which are both in some sense “representational” and both in some sense “singular.” This conceptual homophony between singular content and target is, I believe, a primary reason that the latter has not been clearly distinguished from the former.

At the outset I characterized the depiction literature, in so far as it is comital on the matter, as defining accuracy exclusively in terms of singular and attributive content. To this the Three-Part Model added target. A more nuanced description might be that the literature recognizes a distinction between attributive content and *some* singular element, but conflates aspects of singular content and target. For instance, this singular element is supposed to reflect which particular objects a picture is of (like singular content), but it is also often thought that the same singular element, when combined with attributive content, determines an accuracy value (like target). For these reasons, a more careful statement of the contribution of the Three-Part Model is that, unlike previous accounts, it not only recognizes a role for target (as argued in the previous section), but distinguishes it from the role played by singular content.

This position can be characterized by a pair of sub-theses, each of which is defended separately below. Together they entail Thesis 2 set out above.

Thesis 2A. A picture’s singular content is not determined by its attributive content or target.

Thesis 2B. A picture’s target is not determined by its attributive content or singular content.

Here, saying that one representational relatum is not determined by the others does not imply that it is unconstrained by them. The point for Thesis 2A is to hold that pictures have genuine singular content, in a sense which is not merely derivative in some way of the picture’s target and attributive content. Likewise for Thesis 2B, the point is that pictures have genuine targets, which are not merely derivative of singular and attributive content. If both theses are right, then singular content and target can vary independently of one another.

In what follows I’ll argue for Thesis 2 by appeal to cases which appear to dissociate singular content and target. The Three-Part Model is well-suited to handle such cases, since it treats these as independent relata. I’ll show that theories which fail to treat singular content and target as independent elements cannot explain key semantic facts about these cases.

3.1. Thesis 2A: Independence of Singular Content

I first argue for for the Three-Part Model's Thesis 2A, that pictures have singular contents independent of their attributive contents or targets. I begin with the assumption that pictures are *of* or *about* particular objects. It is in this sense that Picture E is a depiction of *the Sphinx*, or depicts *the Sphinx* as having certain shape and texture properties. As I've suggested, in the Three-Part Model, facts about which objects a picture is *of* are direct reflections of the fact that those objects are constituents of the picture's singular content.

The skeptical alternative is that objects play no role in content, and pictorial content is instead PURELY ATTRIBUTIVE: assertoric pictures have contents and targets, but their contents merely attribute properties to their targets, but they do not express singular content. Defenders of a purely attributive approach to pictorial content need not deny that pictures are *of* or *about* particular objects, but they must offer an alternative method for accounting for these facts. Two strategies in particular present themselves. The first attempts to derive *of*-facts primarily from features of the picture's target scene, while the second attempts to derive such facts primarily from features of the picture's attributive content.

The SYNECDOCHIC STRATEGY trades on the idea that talk of what a picture is *of* is just a way of picking out its target via its *visible parts*.¹⁹ Thus, roughly, to claim that Picture E is a depiction of *the Sphinx* is just to point out that E's target is a particular scene in 1820's Egypt, and the Sphinx is visible in that scene. Schematically, for any picture *P* and object *O*:

Depiction *of*: Synecdochic Strategy

P is of *O* if and only if

- (i) *P* aims at a target scene *S*;
- (ii) *O* is a part of *S* which is visible from the viewpoint of *S*.

The concept of *visibility* here is not quite the familiar one of perceivability, but a geometrical adaptation for which no viewer or optics need be involved. An object *O* is VISIBLE in a scene *S*, whose viewpoint is *V*, just in case, in a geometrical projection of *S* relative to *V*, some parts of *O* are projected onto the picture plane. Thus, if an object is wholly occluded by some surface (relative to *V*) it won't project to the picture plane, so won't be visible. Otherwise, if it is still within the picture frame, it will count as visible.²⁰

19. This is approximately the tack I take in Greenberg (2013: 221) where I write: "Informally, I will often talk of pictures representing objects, but only insofar as those objects are parts of scenes." This is consistent with the fact that I describe pictorial "content" in purely attributive terms, saving "reference" for scenes.

20. The definition of geometrical projection at work here will vary by system of depiction; some systems rely on linear perspective projection, others on isometric projection, and so on

The synecdochic strategy has considerable appeal, for it delivers correct verdicts for large swaths of cases, including all cases in which pictures are fully accurate, and all cases where the inaccuracy is limited to attributive inaccuracy. In this way, perhaps all facts about what objects a picture depicts can be reduced to facts about what target it aims at.

A second approach to a purely attributive view of pictorial content seeks to derive *of*-facts primarily from a picture's attributive content (as opposed to its target). This **DESCRIPTIVIST STRATEGY** was originally outlined and criticized by Lopes (1996: 93-97). The idea is to think of pictures as akin to names on a descriptivist analysis: in the first instance they express sets of properties, and in the second, derivatively pick out individuals which satisfy these properties (Kripke 1972). In particular, this strategy holds that the attributive content of a picture, together with its target, pick out various objects, and these are the objects which the picture is of. Schematically:²¹

Depiction of: Descriptivist Strategy

P is of *O* if and only if

- (i) some region of *P* expresses the attributive content *A*;
- (ii) *P* aims at target scene *S*;
- (iii) *A* and *S* uniquely specify *O*.

In clause (iii), the manner in which the content *A* and the scene *S* uniquely specify *O* may be fleshed out in different ways. But the general idea is easily illustrated by the case of Picture E: E's target is culled from the actual world; the picture's unusual attributive content is satisfied by only one thing in the actual world—the Sphinx; hence, the picture is *of* the Sphinx. This approach differs from the synecdochic strategy by trading a requirement of mere visibility for one of descriptive accuracy. Thus the descriptivist account strengthens the descriptive condition on depiction-of, but allows latitude in the visibility condition.

Yet I hold that neither the synecdochic, nor descriptivist, nor any other purely attributive strategy can successfully account for facts about what pictures are of. This conclusion is based on cases where the picture is clearly *of* some object, but this fact cannot be derived from facts about the target scene. The only way to account for what a picture is of, in such cases, is to independently posit objects in the singular content, as per Thesis 2A.

But cases like this are not trivial to come by. Normal cases of accurate depiction won't do, because these are situations where the objects a picture is of are present and visible in the target. Both the synecdochic and descriptivist (Greenberg 2013; Giardino & Greenberg 2015). As a consequence, what counts as visible will also vary with system of depiction.

21. Clause (i) refers to "some region" of *P*, rather than *P* itself, since a given picture may be of multiple objects, corresponding to different regions of the picture plane.

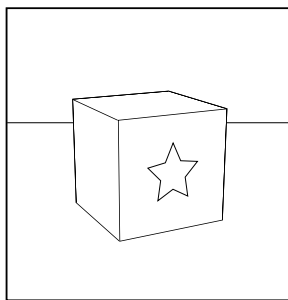
strategy were designed for such situations. In addition, the most familiar forms of inaccurate depiction also fail to make the point. These are cases where the objects a picture is of are visible in the target scene, but the picture misrepresents these objects, say, with respect to color, or shape, or position. In these situations, so long as the misrepresented objects are visible in the target, they will be accounted for by the synecdochic strategy; and so long as the misrepresentation is not too great, they may be explained by the descriptivist strategy as well. Thus, for a great range of cases, the Three-Part Model does not make obviously different predictions than these purely attributive strategies.

Instead, the most compelling argument for the independence of singular content derives from cases where a picture is *of* some object, but that object is not present in the target scene. Such cases effectively dissociate singular content and target to such an extreme degree that attempts to derive one from the other are undermined.

In what follows, I'll describe a case like this, involving an artist who undergoes a hallucination while drawing from life. Note that hallucination itself is inessential to the argument; intent to misrepresent or false background beliefs could deliver the same result. It is also inessential that the picture be drawn from life; drawing from memory or description would work just as well. Here is the case:

Object Hallucination

Yesterday I saw for the first time a particular cube, named Cubey, with a star painted on it, sitting in my garden. (Call the scene defined by this world, time, and viewpoint *Garden*.) Today my desk is empty. I sit down at my desk, and set out to draw what I see. (Call this scene *Empty-Desk*.) But at this point I unwittingly suffer from a partial hallucination, in which it seems to me that Cubey is sitting on the desk before me. I proceed to draw the situation I take myself to be seeing, producing Picture A:



Picture A

The following two facts seem to be implied by my description of the case, and I assume them in what follows:

Assumption 1: Picture A is *of* Cubey.

Assumption 2: Picture A is not accurate.

Assumption 1 corresponds to the intuition that the picture is *of* Cubey, and that it depicts Cubey *as* sitting on the desk. Note that this is *not* the theoretical assumption that Cubey is part of the picture's singular content, only that the picture is *of* Cubey; the stronger claim I will have to argue for. Assumption 2 reflects the fact that although the picture depicts a particular object as located at a certain position, that object is not in fact located at that position— hence the picture must be inaccurate.

The Three-Part Model straightforwardly accommodates these two assumptions. For the first: Picture A is *of* Cubey because Cubey is part of the singular content of the picture. For the second: Picture A is inaccurate because the target of the picture is *Empty-Desk*, and the content of the picture doesn't hold at *Empty-Desk*. In the Three-Part Model, no conceptual tension is created by a picture whose singular content isn't present (or visible) in its target; such content is simply inaccurate. The same cannot be said for purely attributive approaches to pictorial content, as I'll now demonstrate.

To begin, I note that *Empty-Desk* is naturally construed as the target of Picture A. This was the scene before my eyes at the time of drawing, and what I set out to draw. And if *Empty-Desk* is the target of Picture A, that straightforwardly explains the second assumption, that Picture A is inaccurate. There is some temptation to think that the target of A might instead be the non-actual scene which I *took* myself to be seeing, which I in fact rendered accurately. But this is to confuse the content I intended to express with the target I intended to accurately depict. It's true that I was in a perceptual state with a certain content, and I did set out to express that content through my picture; that was my expressive intention, and in this case, the intention was fulfilled. But in addition, I set out to make an accurate depiction of a scene in the world, regardless of whether my perceptual content was accurate at that scene— and in this respect, my intentions were thwarted. It is only the scene in the world which the picture is intended to be accurate at which counts as the target. In this case, it must be *Empty Desk*.²²

The challenge now facing both the synecdochic and descriptivist strategies is to explain how Picture A can be *of* Cubey (Assumption 1) without allowing that Cubey is in Picture A's singular content. Now we can see that the synecdochic strategy is directly counter-exemplified by the Object Hallucination case. For

22. One might wonder if *Garden* is in fact the target of the picture; but there is little to support this, since that is clearly not the scene I set out to draw, even if it is part of the causal source of my hallucination. Indeed, we could modify the case so that Cubey was originally spotted on my desk, rather than my garden, in such a way that Picture A would have been *accurate* at that scene. Then Assumption 2 would rule this out as the target, leaving *Empty-Desk* as the only available option to explain the sense of inaccuracy.

there, the picture is clearly *of* Cubey. But Cubey is not a visible part of *Empty-Desk*, relative to its viewpoint. Thus, in general, what a picture is of cannot be assimilated to the visible parts of its target, and, for all its initial plausibility, the synecdochic strategy quickly unravels.

The descriptivist strategy also fails, for related reasons. If the manner in which the attributive content and target pick out an object is too unrestrictive, then too many objects will satisfy the attributive content of a given picture, and none will be uniquely specified. For example, it is natural to think that an object *O* might be specified as the unique object which satisfies the attributive content *A* in the *time* and *world* of the target scene *S*. (By analogy, as Kripke (1972) construes descriptivism, it holds that a single definite description picks out different objects at different possible worlds.) But as Lopes (1996: 97) has argued, this cannot work.²³ Applied to the case above, Lopes' point is that there may be indefinitely many other indiscernible cubes at the world and time of the picture's target scene, which satisfy the picture's attributive content. Yet Picture *A* is of Cubey only—not the many other possible but spurious cubes.

So some more restrictive approach to picking out the relevant object is required. But the same problem of spurious cubes will continue to arise unless the range of available objects is restricted to those which are *visible in the target of the picture*.²⁴ But then, note, we have just recreated the synecdochic strategy, along with its characteristic defects. For as we saw, such a restriction is too narrow; in the Object Hallucination case, Picture *A* is of Cubey, but Cubey is not visible in the target scene. Thus there is no valid way to use attributive content, even together with a picture's target scene, to specify the objects the picture is of.

Lopes (1996: 96-97) highlights an additional problem facing the descriptivist, stemming from cases in which a picture misrepresents an object by misattributing certain properties to it. Such an object would not satisfy the picture's attributive content in any straightforward way, raising doubts about the descriptivist's basic strategy of picking out objects via attributive content. (Note that the synecdochic strategy does not face the same problem.) The only feasible response is to hold that for a picture to be *of* an object is for that object to satisfy the picture's attributive content *to some limited degree*. But then all the problems highlighted above come back, only with more force. For now one must attempt to uniquely pick out Cubey from all the objects at the world and time of the target scene that fit the attributive content *to some degree*. There seems to be no chance of

23. Lopes (1996: 94) loosely characterizes the descriptivist view as holding that a picture is *of* an object when its attributive content uniquely specifies that object. He doesn't make reference to times or worlds explicit, but these are clearly necessary to the viability of the descriptivist strategy.

24. Though Lopes does not discuss this option, it would answer his original objection to descriptivism, since it adequately distinguishes the accurate, qualitatively identical portraits of indiscernible twins.

picking out Cubey, not only from other cubes, but also from other objects that are sufficiently similar.²⁵

Together, these considerations rule out views according to which what a picture is of can be derived purely from its attributive content, its target, or some combination of these elements. It follows that, for at least some of the objects a picture is of, these objects directly constitute its singular content.²⁶ In the language of Kaplan (1989), we might say that pictures are devices of *direct reference*, for the objects which they are *of* are not merely specified by an intermediary description (e.g. the attributive content), but are themselves parts of the content. Of course, pictures also express attributive content, distinguishing them from standard cases of directly referring terms in language. In this sense, pictorial content is both singular *and* attributive, making pictures both directly referential *and* descriptive.

3.2. Thesis 2B: Independence of Target

I turn now to Thesis 2B: that a picture's target does not depend on that picture's attributive content or singular content.

The alternative view is that pictures simply have no target scenes relative to which they are evaluated for accuracy. I'll call this the INDEX-FREE APPROACH. According to the index-free approach, pictures have attributive and singular contents, but no targets. But since pictures are accurate or inaccurate, content itself must intrinsically determine an accuracy value, without recourse to a further parameter of evaluation. This presents a challenge, since the objects which are normally thought to make up singular content only instantiate their properties and relations relative to a world and time. Here the index-free theorist proposes to build worlds and times *into* the objects that make up the singular content. Singular content, on this account, is not made up of "standard" objects, but world- and time-bound individuals. The idea is that by rooting the singular content of a picture to a particular time and world, ascriptions of attributive content may simply be accurate or inaccurate, obviating the need for an additional index.

To develop this strategy more carefully, I introduce the notion of time- and

25. In addition to the objection outlined here, I have serious misgivings about the descriptivist strategy which echo Kripke's (1972) own objections to descriptivism in the nominal realm. It seems to me that it is possible for a picture to be *of* a given object but misrepresent it so thoroughly that virtually any other (or no) object satisfies the attributive content of the picture. (I have in mind the likes of children's scribbles, or drawings made while blindfolded. See Greenberg (2013: 225) for such an example.) If this is true, there would seem to be no recourse for the descriptivist. Yet my claim that such cases are possible is contentious, and I don't intend to argue the point here. For opposing views, see e.g. Hopkins (1998: 30) and Abell (2009: 212).

26. Alternatively, one might think of singular content as being made up of meaningful parts which intrinsically *determine* particular objects, in the manner of "de re senses" (McDowell 1984). For the sake of simplicity, I'll usually just talk of objects being "in" a picture's singular content, but officially I remain neutral between the constituent and determination formulations.

world-bound objects. Such objects correspond to temporal and modal *slices* or *parts* of ordinary objects. For example, Cubey persists through time in the actual world, and has a variety of exciting careers in other possible worlds and times. But for each such world w and time t we may derive from Cubey a time- and world-bound object C , which exists only at w and t . To fix ideas, let us say that “ $X@wt$ ” denotes that object which (i) exists only at w and t ; (ii) completely overlaps X at w and t ; and (iii) has all the same intrinsic properties as X at w and t . If X does not exist at w and t , then the denoting expression is empty. Let us say that $X@wt$ is the INSTANCE of X at w and t . In addition, since scenes determine both times and worlds, let us say that, for a scene S and object X , “ $X@S$ ” (the instance of X at S) denotes $X@wt$ where w and t are the world and time of S .

The index-free approach proposes that it is *instances* of objects (and not their transworld, transtemporal parents) which make up the singular content of pictures. With this assumption in place, the theorist can now advance a plausible (albeit rough) account of accuracy that is not relativized to a target:

Accuracy: Index-free Approach

P is accurate in c if and only if

all world- and time-bound object instances in P 's singular content in c instantiates the properties in P 's attributive content in c .

Here it makes sense to talk of properties being instantiated without reference to a world or time, because instances of objects, in virtue of existing only at single times and worlds, have their properties absolutely.²⁷

Further, the index-free theorist holds that ordinary judgements about what a picture is *of* can be derived from facts about singular content. A picture is of an object just in case the picture has one of its instances in its singular content:

Depiction of: Index-free Approach

P is of O if and only if

there is some w and t such that
 P has $O@wt$ in its singular content.

Finally, which object instances make up a picture's singular content are thought to be determined by something like a relation of pictorial reference, just as they are in the Three-Part Model. At the very least, for an object instance to be part of

²⁷ In fact, viewpoint-relative relations, legion in pictorial content, still pose a challenge. Such relations are instantiated only relative to a viewpoint, but this appears to be the kind of parameter of evaluation which the index-free theorist sought to avoid. Perhaps there is a way around this. Perhaps the index-free theorist can claim that viewpoints themselves are part of pictures' singular contents. I am skeptical that such a proposal would work, but for the sake of argument, I give the index-free approach the benefit of the doubt.

a picture's singular content, there must be some causal connection between the instance and the creation of the picture, typically mediated by the representational intentions of the artist.

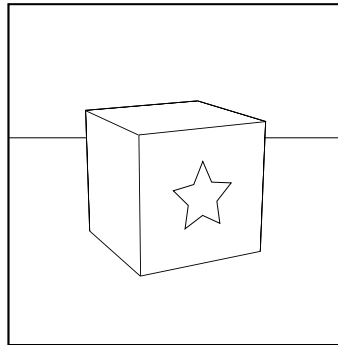
We can now see how the Object Hallucination case might be accounted for under the index-free approach. Recall that the scene before the artist is *Empty-Desk*, but the scene at which the artist originally viewed Cubey, and with which the hallucination is causally linked, was *Garden*. A natural thought is that, because of this link, the region of the picture that depicts Cubey has as its singular content *Cubey@Garden*. Other parts of the picture, for example the regions which depict the desk, will have instances of objects in *Empty-Desk* as their singular contents.

The first assumption about the case is that Picture A is of Cubey. Following the formulation above, the theorist holds that this is so because *Cubey@Garden* is in the picture's singular content. The second assumption is that the picture is inaccurate. This can now be accounted for directly: the picture attributes to *Cubey@Garden* the property of sitting on a desk; but in fact (in *Garden*) *Cubey@Garden* was sitting in a pot of flowers. Thus the picture is not accurate. Thus it seems that the index-free theorist can avoid the challenge posed by the Object Hallucination case.

Although the index-free approach has an answer to the Object Hallucination case, the deeper problem posed by cases *like* it is merely postponed. Ultimately, the index-free approach is committed to a overly restrictive connection between the objects a picture is of, and the world and time relative to which its accuracy is evaluated. Cases which dissociate these elements even more extremely than Object Hallucination therefore challenge its foundational assumption. In this spirit, consider the following:

Scene Hallucination

Yesterday I saw for the first time a particular cube, named Cubey, with a star painted on it, sitting on my desk. (Call the scene defined by this world, time, and viewpoint *Desk*.) Today, I visit my favorite forest. I sit down beneath the trees, and set out to draw what I see. (Call this scene *Forest*.) But at this point I unwittingly suffer from a holistic hallucination, in which the scene I perceived the day before appears as if before me—I no longer perceive the forest. In this confused state, I do not realize I am hallucinating. I proceed to draw the situation I take myself to be seeing, producing Picture B.



Picture B

Here, as before, I assume that two facts follow from my description of the case: first, Picture B is of Cubey; and second, Picture B is not accurate. But now, the index-free strategy of binding objects to worlds and times no longer helps. The origin of the hallucination is the scene *Desk* perceived on the first day, so it is natural to think that the singular content of Picture B includes Cubey@*Desk*. But if so, then it seems the content of Picture B would have to be *accurate*, since, by stipulation, Cubey@*Desk* has all of the properties the picture ascribes to it. But this is the wrong prediction.

Alternatively, perhaps the singular content of Picture 2 includes Cubey@*Forest*. (Recall that “Cubey@*Forest*” refers to Cubey at the time and world of the scene *Forest*, even if, as is the case here, Cubey is not actually visible in that scene.) But we may extend the case by stipulating that by the time of *Forest*, Cubey’s shape and position have not changed since the time of *Desk*. In that case, the attributive content of the picture (that of a cube sitting on a desk) would still be accurate of Cubey@*Forest*. But this again is the wrong prediction. Attempting to further restrict Cubey to objects *visible* in *Forest* yields no gains, since Cubey is not in fact visible in that scene.

The moral here is that invoking time- and world-bound instances of individuals is not enough. Even when *all* of the objects in a picture's singular content are bound to the same scene, it may be that the picture must *still* be evaluated for accuracy at yet another scene. This corresponds to the additional index of evaluation championed by the Three-Part Model.²⁸

The lesson is driven home if we now consider the Three-Part Model's treatment of the Scene Hallucination case. The Three-Part Model holds that the content of Picture B is straightforward— it depicts Cubey (its singular content), as being located in a certain position, sitting on a desk, and so on (its attributive content). What is unusual is that Picture B aims at the target scene *Forest*, rather than *Desk*, and is therefore inaccurate; it is inaccurate both because of the properties it attributes to the scene and because of the objects whose existence it posits.

The Scene Hallucination case shows unambiguously that content cannot determine target. This implies that two pictures with the same content can have different targets. We can extend the Scene Hallucination case to make this point vivid. Compare Picture B, which was drawn at *Forest*, with a new Picture C, qualitatively identical to B, accurately drawn from life in *Desk*. According to the Three-Part Model, B and C have identical contents, both singular and attributive. Nevertheless, B is inaccurate, because its target is *Forest*, while C is accurate because its target is *Desk*. Such comparisons illustrate the basic dimensions of freedom between content and target defined by the Three-Part Model.

In sum, only the Three-Part Model can explain our judgements about cases in which singular content and target dissociate from one another, as in the two cases above. Two-part theories which collapse or reduce these elements lack the resources to account for such variation.

4. Non-Factual Targets

In this section I defend the model by showing how it can handle a range of potential counter-examples. The examples are culled from the phenomena of *counterfactual* and *generic* drawing. They involve pictures that are used in a manner which is intuitively assertoric, but which seem to have no particular or actual

28. In a personal communication, John Kulvicki suggests another tack. Instead of thinking of the singular content of pictures as scene-bound objects, model them instead as $\langle \text{object}, \text{scene} \rangle$ pairs. A picture is accurate when, for each such pair, the object in the pair instantiates its associated properties, at the scene in the pair. Allowing for suitable freedom between the elements of the pair, the key assumptions in the Scene Hallucination case are straightforwardly accounted for. But I count this as a formal variant of the Three-Part Model, for it crucially recognizes the same three representational relata for pictures, and allows that each is independent of the others. The choice between this formalization of the Three-Part Model, and the one I have offered in the text, is a matter of parsimony. My own formalization has the virtue of making the structure of pictorial representation commensurate with that of language.

target. Such examples are *prima facie* challenges to the Three-Part Model's claim that all assertoric pictures have targets. The solution, I propose, is to allow for targets which are, on one hand, not maximally specific, and, on the other, not sampled from the actual world. With this kind of NON-FACTUAL target, the Three-Part Model can be extended to a range of new cases.

I begin with the class of what I will call COUNTERFACTUAL pictures. These are drawings which are specifically produced to illustrate counterfactual scenarios. Suppose I ask you, as my architect, what my house *would* look like if such-and-such modifications were made. If you reply with a drawing, this seems to provide an informative answer to my query, and we may judge it accurate or not, according to whether the house really would look like that under the envisioned conditions. Such pictures may be deemed *assertoric*, for they directly convey information, and they are naturally evaluated for accuracy.²⁹ According to the Three-Part Model, assertoric pictures derive their accuracy values by comparison with a target. But what if anything fulfills the role of target here?

Cases like these present two challenges. The first concerns the modal status of the target. It is clear that, in the case described, the picture is not accurate at any *actual* target; if it is accurate at a target, that target must be counterfactual. So the solution here is to allow that targets need not be culled from the actual world. Of course, one cannot *ostensively* pick out a counterfactual target in the same way one may pick out an actual target in the course of, say, life drawing. But we have already made allowances for descriptive selection of targets— this is what happens when I aim at an (actual) scene which is relayed to me merely via description, as in the case of a police sketch. In the case of counterfactual depiction, the artist's descriptive intentions simply pick out a target culled from a merely possible world.

The operative description must specify the modal relationship between the actual world and the counterfactual target, in order to capture the assertoric force of the original image. It is not enough to say that that the picture is merely accurate at *some* possible world, but rather, that it is accurate at a possible world which is counterfactually related to the actual world in a specific way. Here I appeal to the extensive literature on the semantics of counterfactuals in language (Stalnaker 1968; Lewis 1973; Kratzer 2012). Authors in this tradition have attempted to give precise conditions whereby a counterfactual supposition may pick out a set of worlds that are suitably "nearby" the world of evaluation. We may assume that much the same mechanisms are in play, albeit at the level of thought, when artists'

29. It is important to distinguish architectural *plans* from the kind of drawing here, which (for certain styles of depiction) is sometimes called a *rendering*. The former are plausibly expressions of intention, or instructions to builders, and are not assertoric. The latter are illustrations of established plans, and *are* assertoric, though not typically aimed at the actual world.

intentions pick out a set of possible worlds as their targets in counterfactual depiction.³⁰

The second challenge raised by counterfactual depiction has to do with the *specificity* of the target. Thus far I have proceeded on the assumption that targets are modeled as individual scenes. But my comments above suggest that we may need to reconceive pictorial targets as *sets* of scenes instead. For life drawing, the assumption of a singleton target scene seems reasonable: I intend to draw a bit of the *actual* world, and I intend to draw it from *this* viewpoint. These intentions are sufficient to pick out a particular scene as target.³¹ But for counterfactual representations, it isn't plausible that intentions are so specific. I intend to draw the counterfactual scenario as we discussed it, but our discussions did not determine a *unique* possible world, they specified only a *partial* possible situation. And partial situations like this are modeled by sets of worlds.

If targets are understood as sets of scenes, we must adjust the definition of accuracy accordingly. Here I assume that we maintain the core notion of accuracy at an arbitrary scene, and define accuracy at a *set* of scenes in terms of the former notion:

In a context *c*, *P* is accurate at a set of scenes *S* if and only if there is at least one scene *s* in *S* such that, in *c*, *P* is accurate at *s*.

Recall that for a picture in context to be accurate *simpliciter*, its content must be accurate at its target. Thus, when a picture's target is non-factual, the picture is accurate *simpliciter* when there is *some* scene in its target such that the content is accurate at that scene.

Here it is significant that accuracy at a set requires accuracy at *some* member of the set, rather than accuracy at *every* member. Consider again the hypothetical modification to my house. The resulting picture could be quite specific— showing *one* way the house might look if the modifications were made— even though the preceding discussion may have been comparatively open-ended. (For example, the picture might depict the light on the house coming from a specific angle, even though this was not antecedently specified.) Typically, no one picture will be accurate at *every* scene compatible with a counterfactual description of a scenario,

30. An interesting possibility arises here. *Counterpossible* counterfactuals are those whose antecedents specify impossible situations, as in "if there were a square circle, then..." Whatever the semantics of these counterfactuals, itself an open question, they might be deployed to specify *impossible targets* relative to which pictures with impossible content could be accurate. Presumably this is the only kind of situation in which impossible pictures may be accurate.

31. Even here it is plausible that artist intentions allow some leeway in the particular viewpoint selected. For example, do intentions select the viewpoint located at one or the other of the artist's eyes? Or somewhere in between them? If there is indeterminacy here then the move to model targets as sets may even apply to many cases involving factual targets.

but it may very well be accurate at *some* such scene. Our intuitions of accuracy seem to track the latter condition.

Though it deserves more discussion than I can afford here, I expect that the same basic strategy can be extended to the domain of fictional drawing. Fictions, like counterfactual discourses, can be thought of as specifying sets of possible worlds. Fictional pictures, in turn, can be used to illustrate or extend an existing fiction, or to specify a new one. Though a given scenario may be fictional, there is nothing amiss in evaluating a picture for accuracy *relative* to that fiction (following Lewis 1978). Thus I may draw more and less accurate portraits of Sherlock Holmes. In the present framework, that would mean evaluating the picture for accuracy at a set of scenes compatible with the possible worlds specified by the relevant fiction.

The case of *GENERIC* pictures presents a related and in some ways more demanding challenge for the Three-Part Model.³² Commonly found in textbooks and encyclopedias, generic depictions are pictures which depict an individual (e.g. Obama) or a kind of thing (e.g. eagles), but not at any particular time or place. As it were, such pictures present a “normal view” of the object in question. Generic depictions are clearly assertoric, in the sense that they are meant to convey accurate information about the actual world, but once again, there seems to be no particular time or place at which they are supposed to be accurate. In fact, in some cases, there may *be* no actual time or place at which they are accurate. For example, an encyclopedia illustration of an eagle may be naturally judged accurate even if, by happenstance, no actual eagle was ever positioned so as to perfectly realize that picture’s spatial content.

It is tempting in these cases to eschew the framework of the Three-Part Model for a seemingly simpler treatment. Perhaps the “target” of a generic picture is in fact a particular *object* (in the case of the generic portrait of Obama) or a *kind* of object (in the case of the generic illustration of eagles)—rather than a scene. On this way of thinking, some element of the content is itself the index relative to which a picture’s accuracy is to be measured. There is no role here for the additional involvement of a target, as I have envisioned it.

But this suggestion cannot work, at least as stated. The problem is that a picture cannot be assessed for accuracy simply relative to an individual. Individuals have different properties at different worlds, and at different times within those worlds. Even a generic picture of Obama would be inaccurate relative to a possible Obama with a differently shaped face, or actual Obama as a child. So it is necessary to define at least a world (or set of worlds) and a time (or time span) relative to which the picture will be assessed. But at this point note that a world and time already make up two elements of the target as I conceive it. At the very

32. Thanks to an anonymous reviewer for pressing the challenge posed by generic pictures, and to Susanna Siegel for discussion of a solution.

least, the idea that we might be able to dismiss an index of evaluation, and define accuracy for generic pictures in terms of content alone, cannot work.

How might the Three-Part Model handle the case of generics? Here I'll focus on the case of generic pictures of individuals. I propose that the special difference between generics and other pictures lies in the selection of target, rather than in their content or definition of accuracy. And the targets of generics, like counterfactual depictions, are non-factual: they comprise sets of scenes, none of which need be actual. For generic pictures, however, targets are picked out in a distinctive way, by reference to what is normal for an individual in a part of that individual's life. The idea is that the target of a generic picture is the sets of scenes which contains an individual in a manner which is *normal* for them— where the notion of normalcy, borrowed from the literature on generics in language, is something like “according to expectations,” or “according to an archetype,” with no requirement of actuality or statistical regularity (Nickel 2016). Roughly, the target of a picture relative to an individual *O*, and timespan *T* of *O*'s life, is the set of scenes *S* such that *O* appears as they normally would during *T*, in *S*. Then such a picture is accurate at its target in the standard way, just in case *some* scene in the target is compatible with the picture's content. Consequently, for a generic picture of Obama as an adult, in context, to be accurate, its content must be accurate at some particular possible scene where Obama appears as he normally would during adulthood.³³

Note that, on this account, no particular viewpoint, or set of viewpoints need be picked out *de re* by the artist in selecting a target. The generic picture is accurate if it is accurate at *some* particular scene in the target, but the targets here may vary arbitrarily with respect to viewpoint. (Still, there is reason to believe that only certain kinds of “standard” viewpoints are appropriate for generic depiction.³⁴ A drawing of a seated Eagle from *underneath*, for example, would plausibly *not* count as a generic picture of an eagle, except perhaps in a specialized context.)

The account of generic pictures we've arrived at has the same general form as that offered for the analysis of counterfactual depiction, relying as it does on the notion of a non-factual target, and the accompanying definition of accuracy. Although the phenomena of counterfactual and generic depiction were initially presented as challenges to the Three-Part Model, ultimately, I think the considerations here reveal the flexibility of the model to illuminate a wide range of cases—

33. The case of generic *kind* pictures is analogous. Roughly, the target of a picture relative to a kind *K* is the set of scenes *S* such that a normal *K* is visible in *S*. A generic picture of an eagle is accurate because it is accurate at a scene where a normal eagle is visible. (Perhaps a temporal restriction is necessary here as well.) At any rate, this a first-pass analysis of this kind of depiction.

34. See Huffman (1971: 298) for the concept of a “general” viewpoint, and Willats (1997: 23-24) for discussion. Hagen (1986: ch. 6) describes a number of viewpoint “standards” in non-Western depiction.

everything from photography and life drawing to counterfactual and generic depiction.

5. Conclusion

In this essay I've argued for the Three-Part Model, focusing exclusively on the case of pictorial representation. But the considerations raised suggest that the same basic semantic architecture may arise in representational systems beyond depiction. As I noted in Section 1, the key elements of (i) singular content, (ii) predicative content, and (iii) an independent index of evaluation are already incorporated in standard accounts of sentential truth within philosophy of language and linguistics. These ingredients bear clear similarities to the notions of singular content, attributive content, and target in the Three-Part Model.

Analogous remarks also apply to the cases of visual perception and mental imagery. In the philosophy of perception, it is already commonly thought that perceptual states have both singular and attributive content (e.g. Burge (2010); Siegel (2011)). Following the spirit of Cummins (1996), I suspect the notion of target also finds a home here. Since vision arguably functions to represent the environment immediately surrounding the perceiver, it is natural to think that perceptual states have targets fixed by the world and viewpoint of the perceiver, at the time of perception. (In this respect, the targets of perceptual states are determined in a manner analogous to those of photography, where the target is always the world and viewpoint of the camera.) Then a perceptual state would be accurate if its singular content instantiates its attributive content relative to the target scene—that is, the world and viewpoint of the perceiver.

By contrast, mental imagery can function to describe both past, present, actual and counterfactual scenarios: I can image what has happened, what will happen, and what would happen. Thus, in the case of mental imagery, it is natural to think that the target may range across worlds and times, just as they do for drawing. A mental image is accurate when it has a target and its singular content instantiates its attributive content relative to that target. Like non-factual depiction, the target of a mental image may often be specified only in counterfactual relation to the actual world.

It remains to be seen, of course, whether the arguments of this essay which stem from the Object and Scene Hallucination cases can be carried over to the domains of visual perception and mental imagery. Supposing they can, however, a distinction between content and target there would not only be an available theoretical options, but a necessity.

If these speculations are on track, then, across a range of modalities, we find the same basic elements that make-up the Three-Part Model: a representation, whose function it is to be accurate, expresses both singular and attributive (or

predicative) content; and this content is evaluated for accuracy (or truth) at a target index. Although the rules that associate representations with contents seem to vary dramatically from depiction, to language, to visual perception and mental imagery, I conjecture that the same basic semantic architecture is characteristic of all sufficiently complex representational systems.

Acknowledgements

This essay has grown from years of conversation with friends, colleagues, teachers, and students, to all of whom I am truly grateful. Special thanks to Josh Armstrong, Sam Cumming, Katie Elliott, John Kulvicki, Sun-Joo Shin, and several anonymous referees for comments on the current draft, and to audiences at the 2016 Central APA, Arizona State University, University of Texas at Austin, and Washington University in St. Louis.

References

- Abell, C. (2009). Canny Resemblance. *The Philosophical Review*, 118(2):183–223.
- Abusch, D. (2015). Possible worlds semantics for pictures. Online manuscript.
- Austin, J. L. (1950). Truth. *Proceedings of the Aristotelian Society, Supplementary Volumes*, 24:111–128.
- Barwise, J. and Etchemendy, J. (1987). *The liar: An essay on truth and circularity*. Oxford University Press.
- Blumson, B. (2009). Pictures, perspective and possibility. *Philosophical Studies*, 149(2):135–151.
- Blumson, B. (2014). *Resemblance and Representation: An Essay in the Philosophy of Pictures*. Open Book Publishers.
- Burge, T. (1991). Vision and intentional content. In Lepore, E. and van Gulick, R., editors, *John Searle and his Critics*, pages 195–213. Blackwell.
- Burge, T. (2010). *Origins of Objectivity*. Oxford University Press, Oxford.
- Burge, T. (2014). Reply to Rescorla and Peacocke: Perceptual content in light of perceptual constancies and biological constraints. *Philosophy and Phenomenological Research*, 88(2):485–501.
- Camp, E. (2007). Thinking with maps. *Philosophical Perspectives*, 21(1):145–182.
- Casati, R. and Varzi, A. C. (1999). *Parts and Places: The Structures of Spatial Representation*. MIT Press.
- Cummins, R. (1996). *Representations, Targets, and Attitudes*. MIT Press, Cambridge, MA.
- Eaton, M. (1980). Truth in pictures. *The Journal of Aesthetics and Art Criticism*, 39(1):15–26.

- Giardino, V. and Greenberg, G. (2015). Varieties of iconicity. *Review of Philosophy and Psychology*, 6(1):1–25.
- Goodman, N. (1968). *Languages of Art: An Approach to a Theory of Symbols*. Bobbs-Merrill, Indianapolis.
- Greenberg, G. (2013). Beyond resemblance. *Philosophical Review*, 122(2):215–287.
- Hagen, M. A. (1986). *Varieties of Realism: Geometries of Representational Art*. Cambridge University Press, Cambridge.
- Hopkins, R. (1998). *Picture, Image and Experience: A Philosophical Inquiry*. Cambridge University Press.
- Huffman, D. (1971). Impossible objects as nonsense sentences. *Machine intelligence*, 6:295–323.
- Hyman, J. (2006). *The Objective Eye: Color, Form, and Reality in the Theory of Art*. University of Chicago Press.
- Hyman, J. (2012). Depiction. *Royal Institute of Philosophy Supplement*, 71:129–150.
- Kaplan, D. (1968). Quantifying in. *Synthese*, 19(1-2):178–214.
- Kaplan, D. (1989). Demonstratives: An essay on the semantics, logic, metaphysics, and epistemology of demonstratives and other indexicals. In Almog, J., Perry, J., and Wettstein, H., editors, *Themes from Kaplan*, pages 481–563. Oxford University Press.
- Kjørup, S. (1974). George inness and the battle at hastings, or doing things with pictures. *The Monist*, 58(2):216–235.
- Kjørup, S. (1978). Pictorial speech acts. *Erkenntnis*, 12(1):55–71.
- Korsmeyer, C. (1985). Pictorial assertion. *The Journal of aesthetics and art criticism*, 43(3):257–265.
- Kratzer, A. (2012). *Modals and conditionals: New and revised perspectives*, volume 36. Oxford University Press.
- Kratzer, A. (2017). Situations in natural language semantics. *Stanford Encyclopedia of Philosophy*.
- Kripke, S. A. (1972). Naming and necessity. In Davidson, D. and Harman, G., editors, *Semantics of natural language*, pages 253–355. Springer.
- Kulvicki, J. (2006). *On Images: Their Structure and Content*. Clarendon, Oxford.
- Lewis, D. (1973). *Counterfactuals*. Blackwell Publishers.
- Lewis, D. (1978). Truth in fiction. *American Philosophical Quarterly*, 15(1):37–46.
- Lewis, D. (1980). Index, context, and content. In Kanger, S. and Ohman, S., editors, *Philosophy and Grammar*, pages 79–100. D. Reidel, Dordrecht.
- Lopes, D. (1996). *Understanding Pictures*. Oxford University Press, Oxford.
- MacFarlane, J. (2009). Nonindexical contextualism. *Synthese*, 166(2):231–250.
- McDowell, J. (1984). De re senses. *The Philosophical Quarterly (1950-)*, 34(136):283–294.
- Nanay, B. (2018). Threefoldness. *Philosophical Studies*, 175(1):163–182.
- Neander, K. (1987). Pictorial representation: A matter of resemblance. *The British*

- Journal of Aesthetics*, 27(3):213–226.
- Newall, M. (2011). *What is a Picture?: Depiction, Realism, Abstraction*. Palgrave Macmillan.
- Nickel, B. (2016). *Between logic and the world: An integrated theory of generics*. Oxford University Press.
- Novitz, D. (1975). Picturing. *Journal of Aesthetics and Art Criticism*, 34(2):145–155.
- Peacocke, C. (1987). Depiction. *The Philosophical Review*, 96(3):383–410.
- Peacocke, C. (1992). *A Study of Concepts*. The MIT Press.
- Recanati, F. (2001). Open quotation. *Mind*, 110(439):637–687.
- Rescorla, M. (2009). Predication and cartographic representation. *Synthese*, 169(1):175–200.
- Ross, J. (1997). *The Semantics of Media*. Kluwer Academic Publishers.
- Schier, F. (1986). *Deeper Into Pictures: An Essay on Pictorial Representation*. Cambridge University Press.
- Siegel, S. (2011). *The Contents of Visual Experience*. Oxford University Press.
- Stalnaker, R. (1968). A theory of conditionals. In Harper, W. L., Stalnaker, R., and Pearce, G., editors, *Ifs. Conditionals, Belief, Decision, Chance, and Time*, pages 41–55. Dordrecht, D. Reidel Publishing Company.
- Willats, J. (1997). *Art and Representation: New Principles in the Analysis of Pictures*. Princeton University Press, Princeton.
- Wollheim, R. (1987). *Painting as an Art*. Thames and Hudson.