Content and Target in Pictorial Representation

Gabriel Greenberg   April 17, 2017   Department of Philosophy, UCLA   [DRAFT]

This paper argues for a particular model of pictorial representation. 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*. A picture is accurate when the content it expresses fits the target scene it aims at. In addition, content in this model has 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 attributive content, singular content, and target. Previous work on depiction has recognized a distinction between singular and attributive content, but it has failed to recognize the semantic role of target. By contrast, I'll argue that the concept of target is essential to correctly describe a range of central cases.

In Section 1, I sketch the core features of the Three-Part Model and situate it in relation to the literature on depiction. In Section 2, I argue from cases that we must accept the distinctive claim of this model, that pictures have independent singular contents and targets. Section 3 shows how the Three-Part Model may be adapted to handle the phenomena of counterfactual and generic depiction. In Section 4, the conclusion, I suggest that the same three-part semantic architectural extends to language, vision, and imagery.

1   The Three-Part Model

In this section, I'll lay out the basic tenets of Three-Part Model and sketch the factors which motivate it. I'll start by outlining the model's core commitments, contrasting it with extant theories of depiction, and indicating the central intuitions which give it impetus. I then go on to spell out the more specific claims of the Three-Part Model, with attention to each of its components parts: picture, singular content, attributive content, and target. My intention throughout this section is merely to present the model; only in the next section will I turn to argue for its key claims.

To begin, 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.

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 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 what I will term SEMANTIC aspects of a picture, because they have to do with how the picture depicts the world. They are reflections of the picture's CONTENT. The situation depicted by a picture in this way is presumed to be *content*, and not a mere informational trace of the world, because it determines a set of substantive accuracy conditions for the picture. When these conditions are not met, the picture is not accurate: it misrepresents. If it turned out, for example, that the Sphinx did not have the shape which Picture E attributes to it, the content of the picture would be no different, but in that case it would be inaccurate. Thus the content of a picture corresponds to how the picture depicts the world— independent of whether the world fits this construal.

Over the last fifty years, scholars of depiction have distinguished between two aspects of picto-

rial content, which, for the sake of standardization, I'll call SINGULAR CONTENT and ATTRIBUTIVE CONTENT. 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.[1] 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*. 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. A natural elaboration of this view makes accuracy a function of both singular content and attributive content: a picture is accurate when its singular content actually instantiates the properties and relations ascribed by its attributive content.[2]

These two aspects of pictorial representation have gone by many names.[3] Goodman (1968, pp. 27-8) distinguishes the *kind* of a picture from its *denotation*. Kaplan (1968, pp. 197-8) differentiates between a picture's *descriptive content* and its *genetic character*. Schier (1986) distinguishes *iconic prediction* from *iconic reference*. Hyman (2012, p. 136) separates out a picture's *sense* and its *reference*. And Greenberg (2013, p. 222) marks a distinction between a picture's *content* and its *referent*. Though the details of the distinction vary from author to author, the emerging consensus is that pictorial representation can be factored, perhaps exhaustively, into singular and attributive components. This factorization is presumed to capture the central function of depiction: for a picture to *depict* is for it to ascribe its attributive content to the objects in its singular content (Kjørup 1978, pp. 62-64); and for a picture to be *accurate* is for its attributive content to be realized by the objects in its singular content (Goodman 1968, p. 38).

I propose that this view of pictorial representation is correct as far as it goes, but ultimately incomplete, to be superseded by the Three-Part Model. Pictures *do* have singular and attributive content, and for a picture to express its content *is* for it to ascribe its attributive content to its singular content. But this characterization does not exhaust the nature of depiction and is insufficient to define pictorial accuracy. The distinctive claim of the Three-Part Model is that, in addition to expressing singular and attributive contents, pictures have what I call TARGETS. The concept of *target* here is adapted from Cummins' (1996) theory of mental representation.[4] In my construal,

---

[1]It likely includes other elements of the depicted landscape, such as individual dunes and ridges, as well as *parts* of the objects already mentioned— parts of the Sphinx, pyramids, dunes, and so on. The full extent of singular content is an issue I postpone for further inquiry.

[2]Scholars in the depiction literature vary in how explicitly they endorses this claim; for the sake of exposition, I will treat it as a core commitment of the received view.

[3]Note that while most author's writing on the topic have recognized something like singular content, they have not always treated it as part of *content*. Instead, "content" (or cognates, like "sense") are sometimes reserved for what I am calling *attributive* content, while an independent semantic relation ("denotation", "reference") is posited for the expression of what I am calling *singular* content. Indeed, some authors' characterization of singular content has more in common with my construal of *target*. The contrast between the Three-Part Model and the literature is stated more carefully at the beginning of Section 2.

[4]While the ideas here are directly inspired by Cummins presentation, they also diverge from his in important ways. Cummins' notion of a target seems to be that of the *content* which a computational system is supposed to express in context;

the target of a picture is, roughly, the particular spatio-temporal location, anchored at a particular viewpoint, which it is the picture's *purpose* or *function* to be accurate at. In contrast with the received view, the Three-Part Model holds that depiction in fact involves two representational relations; a picture not only expresses content (both singular and attributive), it also aims at a target. An artist setting out to depict the world not only intends to create a picture that expresses a certain content, she also intends that this content will be accurate a particular target.[5]

This difference between the received view and the Three-Part Model, about the nature of depiction, is directly reflected in their definitions of accuracy. On the received view, for a picture to be accurate, its singular content must realize its attributive content. On the Three-Part Model, singular and attributive content alone are insufficient to determine an accuracy value. Instead, for a picture to be accurate, its singular content must realize its attributive content, *in the target scene*. In effect, the *target* is the subject matter of a picture; the *content* of the picture provides a kind of comment; it is the function of the picture to offer that content as a comment on the target. Only, in Three-Part Model, the "comment" involves both attributive and singular elements.

By way of illustration, consider again the case of Picture E. The content of E is apparent— it describes a certain pictorial 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 which the particular spatio-temporal location (in 1820's Egypt) which it targets happens to instantiate the singular and attributive content which it expresses. Had (visible features of) the target been otherwise, the picture would have been inaccurate.

Though content and target are distinct, they can be easily confused, because pictures, on this view, now have *two* representational relata both of which are in some sense "singular"— singular content on one hand, and target on the other. Indeed, according to the Three-Part Model, talking about "what a picture is *of*" is actually ambiguous. A picture is said to be *of* the objects in its *singular content*; this is the sense in which Picture E is *of* the Sphinx. But it is also possible to describe a picture as being *of* a scene by way of picking out its *target*; it is in this sense that Picture E is *of* the Giza Plateau at a certain moment in the 1820's. This conceptual homophony between singular content and target is, I believe, the primary reason that the latter has not been clearly distinguished from the former. (Note that, in this paper, I normally use "depiction *of*" in the singular content sense.)

---

by contrast, my notion is that of the *index* relative to which content is supposed to be evaluated for accuracy.

[5] One might still wonder how the proposal here interacts with that strand of the philosophical literature which seeks to give necessary and sufficient conditions on "depiction." But I find such talk obscure, because "what a picture depicts" is ambiguous: it can variously refer to singular content, attributive content, content in general, or target. My own method is to use the term "depiction" loosely, identifying and debating more specific representational relations, in the way that I have, as required by the phenomena at hand.

The triadic structure of the Three-Part Model reflects common themes in the philosophy of depiction, of mind, and of language. On one hand, the initial distinction between singular and attributive content derives from the traditional literature on depiction and is reinforced by recent work in perception, such as that of Burge (2010) and Siegel (2011). In the linguistic domain, it has long been widely assumed that linguistic content contains attributive elements— these are, if nothing else, the denotations of predicates. The work of Kripke (1972) and Kaplan (1989), among others established that (because of the failure of descriptivism) sentential content typically includes singular elements as well.

On the other hand, the distinction between content and target is nominally due to Cummins (1996), but it is more directly traced to contemporary philosophy of language. There, any number of authors distinguish the *proposition* expressed by a sentence ($\approx$ content) 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. Indeed, the Three-Part Model is the product of attempting to secure, for the visual domain, the flexibility and distinctions which have motivated these developments in philosophy of language. While the Three-Part Model is defended here as an account of *pictorial* representation, the parallels noted above suggest that the same basic architecture may be endemic to forms of representation well beyond depiction, including perception and language. I return to this theme in the conclusion.

A central motivation for the distinction between content and target ushered in by the Three-Part Model is that what content a picture expresses, and the circumstance relative to which it is evaluated for accuracy appear to vary independently. (Such considerations are the basis of the argument for the Three-Part Model presented in Section 2.) Thus 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 have they have the same target— they aim at the same location, time, and viewpoint in Giza, before the Sphinx. One is accurate because its content fits this target, and the other inaccurate, because its content and does not fit with the same target.

In addition, two pictures may also have the *same* content, but *different* targets. For example, suppose one person drew Picture E_1 from life at time 1, and another person drew Picture E_2 from life at time 2. Supposing that both E_1 and E_2 are qualitatively identical to the original E, and supposing that all the same individuals appeared in the scene at time 1 and time 2 (the same dunes, the same rider, the same horse), then E_1 and E_2 may have the exact same content. In the Three-Part Model, the sense in which E_1 and E_2 are representations of *different* times is not be traced to content; rather, it is solely matter of target: the target of E_2 is located at a time after the

time of the target of E_1.[6]

Further motivation for distinguishing contents from targets, construed as particular viewpoint-centered scenes, is brought out by again considering what determines whether a given picture is *accurate*. The received view holds that for a picture to be accurate is for the objects in its singular content to satisfy the properties and relations in its attributive content. But this claim cannot be quite right as stated. In the first place, according to a standard possible-worlds framework, objects do not have their properties absolutely, but only relative to a possible world. Since the Sphinx could have had a different size and shape then it actually does, its properties vary by world. Thus pictorial accuracy must be relativized at least to world, for the properties it attributes to the objects in its singular content are only instantiated at some worlds, and not others. (Nor are pictures always evaluated for accuracy at the actual world; as I'll discuss, hypothetical pictures may perfectly well target counterfactual scenes as well.)

The same types of considerations carry over to times and viewpoints. What properties an object instantiates varies by time; the Sphinx had a nose at one time, and not at another. Thus the accuracy of Picture E's content can only be assessed relative to a time. In the same way, the distinctively perspectival properties and relations attributed by a picture are only instantiated relative to a viewpoint. For example, the Sphinx occludes the Pyramid of Khafre (and not Khufu) relative to some viewpoints; the reverse is true relative to others. Since perspectival relations like distance, visual angle, relative position, and occlusion are expressed by pictures, their contents again can only be evaluated for accuracy relative to a viewpoint. Putting these elements together— world, time, and viewpoint— one arrives at the notion target which I have in mind: a target is a particular location in space and time, centered at a viewpoint.

Thus, according to the Three-Part Model, pictures not only express singular and attributive content, as previous literature on depiction has recognized, they are also accurate or inaccurate at particular viewpoint-centered targets.

Having laid out the core intuitive motivations for the Three-Part Model I now turn to provide a more systematic summary of its component parts, including pictures, singular content, attributive content, and target. To begin, PICTURES, as I shall use the term, are a species of public signal whose form and content are in some sense distinctively visual. Canonical examples include life drawings, engineering plans, architectural renderings, courtroom sketches, medical illustrations, and photographs. Pictures include many works of visual *art*, but are not limited to them. Officially, I will treat pictures as abstract types, specified in terms of their 2-dimensional geometrical and chromatic organization. Particular picture tokens correspond to the instantiation of pictures in

---

[6]Beyond the core tenets of the Three-Part Model, these facts reflect a commitment to "viewpoint neutrality"— the position that pictorial content doesn't itself contain a specific central viewpoint, but is compatible with different viewpoints. I assume viewpoint neutrality throughout; one motivation, highlighted just below, is that it makes room for the intuition that the accuracy value of the content of a given picture may vary when evaluated at different viewpoints.

context, though for convenience, I'll often call these "pictures" as well.

The Three-Part Model is not intended as an exhaustive account of the pictorial representation, and there are several respects in which its scope is limited. First, it is intended to apply only to what I term ASSERTORIC PICTURES: pictures which function to be accurate depictions of the world. Intuitively, such pictures provide information about the way things are; they include drawings from life and memory, courtroom sketches, maps, and photographs. By contrast, IMPERATIVAL PICTURES, like ikea instructional diagrams, architectural plans, or road-side warning signs, function to convey instructions or plans, but not to be accurate; hence they lack targets.[7] A wide range of pictures besides also lack targets because they do not aim at accuracy; these include many doodles, decorative designs, fantasy images, and art works. While nearly all such pictures have content, among them only assertoric pictures have targets, for only assertoric pictures are designed to be accurate. It is assertoric pictorial representation which is the subject of the Three-Part Model.[8]

In addition, even among assertoric pictures, I set aside hypothesized aspects of content that are not directly reflected in accuracy conditions (understood as sets of centered-worlds), such as modes of presentation or other hyper-intensional meanings. Peacocke (1992) and Burge (1991; 2014), among others, have argued for such elements in visual perception. Whether these conclusions carry over to depiction, I leave to future investigation, though I suspect they will.

With these caveats in mind, the schematic relations between singular content, attributive content, and target, as envisioned by the Three-Part Model, are illustrated below:
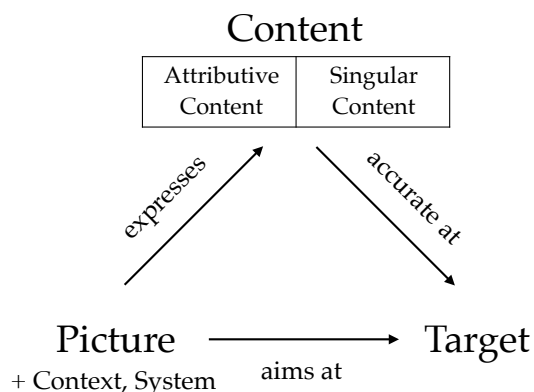
The Three-Part Model

---

[7]See Kjørup (1978, pp. 64-66) and Korsmeyer (1985) for classifications of pictorial speech acts in this vein.

[8]In the background here is an assumption of Frege's classical distinction between *force* and *content* (Frege 1879; Geach 1965). As I conceive it, pictorial content is neutral with respect to force. The same spatial content can be used assertoricaly, to describe the world, or imperatively, to issue an instruction. Hence, when I make reference to the "accuracy conditions" determined by a picture's content, I mean this to be neutral as to whether the picture functions to be accurate or not. A fuller development of the theory of communicative acts for pictures lies outside the scope of this paper.

The CONTENT of a picture corresponds roughly to *what's happening in it* or *what it depicts.* A picture's content, in this sense, determines the conditions under which the picture is accurate, just as the contents of sentences determine truth conditions. I'll say that a picture EXPRESSES its content. Expression, in turn, is relativized to both a SYSTEM OF DEPICTION and a CONTEXT. Systems of depiction are the pictorial analogues of languages; they play the central role of associating the geometrical and chromatic surface features of pictures with elements of attributive content.[9] By contrast, the context of a token picture, understood as the causal and psychological situation in which the picture is created, functions to associate the picture with elements of singular content. But since neither system nor context will play a central role in the discussion to follow, I'll generally suppress reference to them.

While pictorial content is reflected in facts about what objects a picture is of, and what properties it depicts these objects as having, no single such fact captures the full content of a picture. A picture never merely depicts a plant as having green leaves, for instance; it also depicts it as having a certain shape, orientation, location, all relative to a general viewpoint. Thus it is natural to conceive of the contents of a picture as a PICTORIAL SPACE centered at a general viewpoint: a three-dimensional region populated with individuals, properties, and relations, where the locations of these elements are always specified relative to a central perspective position. (Howell 1974; Koenderink and Doorn 2003) Intuitively, a space like this may be accurate or inaccurate at a target scene to the extent that the scene is a realization of that space.

In characterizing pictorial space, I invoke the concept of a GENERAL viewpoint, or viewpoint *type*, in contradistinction with the notion of a PARTICULAR viewpoint, or viewpoint *token*. A picture depicts objects and properties relative to a *general* viewpoint, because the picture's content need not include any *particular* token viewpoint. At the same time, a given pictorial space may be *realized* by some particular viewpoint in a particular space. So while a pictorial content is general with respect to viewpoint, it is typically accurate or inaccurate relative to a particular viewpoint.[10] In what follows, I normally reserve the term "viewpoint" for particular token viewpoints, disambiguating when necessary.

In general, on the Three-Part Model, content is neutral with respect to any particular world, time, or viewpoint. Yes, pictures express pictorial spaces, and pictorial spaces are accurate *at* particular worlds, times, and viewpoints. But such spaces themselves contains no particular information about what world, time, or viewpoint they are to be evaluated at. Thus, what's depicted by a given picture— its content— may be accurate at some worlds and not at others; accurate at some times and not at others; and even accurate relative to some viewpoints, and not relative to others. If the content itself contained these elements, such evaluations would be impossible.

For many purposes, it is useful to model pictorial content within the framework of possible-

---

[9]See Giardino and Greenberg (2015) for an introduction to and defense of the relevant notion of systems of depiction.

[10]In Section 3, I discuss cases in which accuracy may be evaluated relative to a set of viewpoints, rather than a single particular.

world semantics. Such formalism isn't essential to the Three-Part Model, but helps to fix ideas. To this end, I introduce the concept of a "viewpoint-centered world" or SCENE, defined as the pair of a possible world and a (particular) viewpoint. A VIEWPOINT here is conceived of as a particular oriented location in space and time, and carries no implication of a real or metaphorical viewer. Then, following Ross (1997), Blumson (2009), and Greenberg (2016) we may model a picture's content as a *set* of scenes. Formally, such a set may be thought as a relation between worlds and viewpoints: the content of a picture imposes a simultaneous constraint on the way the world is relative to a viewpoint, and the location of that viewpoint relative to the world.

This model of a picture's content fits smoothly with the characterization of a picture's content as a pictorial space centered on a general viewpoint, populated with objects and properties. This is because every such space determines a set of scenes— those pairs of worlds and particular viewpoints for which, when that particular viewpoint occupies the center of the pictorial space, and the world is an instantiation of that space. Of course, an arbitrarily selected set of scenes won't correspond to a pictorial space in this way, as its members may be entirely disjoint. But those sets of scenes which *are* the contents of pictures will all coincide precisely insofar as the picture's represented space is determinate, and diverge insofar as it is indeterminate.

As advertised, pictorial content can be further subdivided into singular and attributive aspects. SINGULAR CONTENT is that aspect that is composed of (or determines) the particular individuals which a picture is *of*. ATTRIBUTIVE CONTENT is that aspect of a picture's content that is composed of (or intrinsically determines) an array of properties and relations. Taken as a whole, a picture's content ascribes the properties and relations in its attributive content to the objects in its singular content. In terms of pictorial spaces, the objects that populate a given pictorial space belong to a picture's singular content, while the properties and relations located in that space, along with the structure of the space itself, belong to its attributive content.

In fact, singular content belongs to the more general category of OBJECTUAL CONTENT; such content may be an INDEFINITE, as is the case when a picture merely depicts *some* cube as being located in a given direction— or it may be SINGULAR, as when a picture depicts a *particular* cube as being located in a given direction. While indefinite objectual content is both theoretically important and common in depiction, it is the treatment of singular objectual content which is distinctive of the Three-Part Model, and it will be the focus of the discussion to follow.

Crucially, in the Three-Part Model, singular content is independent of the other two representational relata, in the sense that it cannot be derived from facts about attributive content or target alone. (I'll argue for this conclusion in the next section.) Since the objects picked out by singular content cannot be derived from other aspects of content, it is natural to conceive of these objects as being genuine constituents of the content itself.[11] In the language of Kaplan (1989), we might

---

[11]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

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, setting them apart 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.

Theorists of depiction have often assumed that the relationship between a picture and the objects in its singular content is one of *pictorial reference*, analogous in important respects to reference in language and thought. (See, e.g. Schier 1986, Lopes 1996.) Whether or not these amount to the *same* relation, there are clear commonalities. As I've just suggested, pictures seem to pick out the objects in their singular content in a manner comparable to directly referring terms in language. And as Kaplan (1968, pp. 197-8) and Lopes (1996, ch. 5) have argued, pictorial singular content is largely fixed by the history and circumstances under which the picture was created, in a manner analogous to that envisioned by the causal theory of reference for names.

I turn next to the concept of relative accuracy, or accuracy *at* a scene. The content of a picture can be evaluated as ACCURATE AT or INACCURATE AT a scene— much the way linguistic propositions are thought to be evaluable as true or false relative to a world or index of evaluation. Since contents are modeled as *sets* of scenes, a content is accurate at a scene if and only if the scene is a *member of* that content. (Note that it is *contents* which are accurate at scenes; in a derivative sense, we may speak of a *picture*, relative to a system and context, being accurate at a scene, when its content is.) In the sense of relative accuracy I have in mind, all pictures with contents have accuracy conditions, even if only some such pictures function to be accurate (hence have targets). Imperatival pictures, for example, do not function to be accurate; nevertheless, they have the same kind of content and accuracy conditions as any other picture.

A picture's content is accurate 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, accuracy at a scene requires (i) that the objects in the singular content actually exist in the scene; and (ii) that the objects so depicted actually 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: the content of a picture corresponds to the set of scenes relative to which that content is perfectly accurate. 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.[12]

---

objects being "in" a picture's singular content, but officially I remain neutral between the constituent and determination formulations.

[12]In this characterization we risk loosing sight of that aspect of pictorial content which, relative to a scene, determines

Accuracy here is meant to be the pictorial analogue of truth for sentences. A picture's content is accurate at a scene when, to the extent that it represents things as being a certain way in that scene, that is the way those things are. Accuracy, in this sense, does not imply realism or closeness to reality. A black and white drawing and a color painting may each be perfectly accurate, albeit relative to different systems, though the experiences they elicit obviously differ in the proximity to normal perceptual experience. A full scale, working model might be closer to reality, in some sense, than a technical drawing, but both may be perfectly accurate. What matters for accuracy is the absence of misrepresentation— not the quantity or type of information represented.[13] This notion of accuracy is not the exclusive denotation of the word "accuracy" in colloquial English, but it is arguably one of them.

Thus far I've discussed the ways in which a picture's content 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 is *intended* or *functions* to be accurate. I'll say that a picture AIMS AT its target scene. 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 succeeds; 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. It is natural to hold that a picture, in context, is simply *accurate*, in an absolute sense, when it is accurate at its target.

So the target of a picture is the scene relative to which it is the picture's function to be accurate or inaccurate.[14] The notion of *function* here is broad. For many kinds of pictures, it is fixed by the artist's intentions or purposes.[15] In life drawing, for example, the target is the scene before the artist (that is, the world and viewpoint of the artist at the time of drawing), because that is what she intends to represent. In drawing from memory, some previously viewed scene becomes the target; in drawing from a description, or drawing a proposed building, the target scene may never have been viewed by the artist. 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

intermediate degrees of accuracy. Still, it should be noted that those scenes at which the content of a picture is perfectly accurate are, intuitively, exactly those which reflect that 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 pictorial semantics.

[13]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 carries content. (See Rescorla (2008) and Camp (2007) for discussion.) The requirement for a picture's accuracy is that its *content* not misrepresent its target.

[14]The concepts of representational function and success here are based on Burge's (2010, pp. 308-315) discussion of representational function in perception.

[15]Where representational intentions are confused (in the style of examples discussed by Kripke (1977)), there may be than more than one scene at which the artist intends her picture to be accurate. Such cases may require the recognition of two targets, or none; I leave this matter to further investigation.

artist, which fix the target.[16] Typically, in these cases, the target is the scene before the lens of a camera. In Section 3, I'll discuss further cases in which picture aim at *counterfactual* and *generic* targets.

For a scene to be a picture's target, it need not be the case that the picture functions to be *accurate* at that scene, merely that it functions to be accurate *or* inaccurate. A pictorial liar may intend to misrepresent my office as larger than than it actually is. This makes my office the target of the picture, even if the artist's goals are satisfied only when the picture is inaccurate. Instead, it seems that it is a norm for sincerity in pictorial communication— but not a condition on pictorial assertion— that a picture's content be accurate at its target.[17]

We can now see why only assertoric pictures have targets, for only assertoric pictures function to be accurate in this sense. By contrast, an imperatival picture, like a standard Ikea diagram, simply issues an instruction; certain constructions may or may not *satisfy* the picture, but the picture itself is neither accurate nor inaccurate. Other kinds of images are neither assertoric nor imperatival; their central function is to please, inspire, stir the imagination, or trick the perceptual system. Examples include many doodles, some kinds of artworks and childrens' drawings, many optical illusions, and most decorative images, like wallpaper designs or fabric prints. Such images, when they are representational, do express spatial content, but the there is no actual or possible scene which the artist in these cases has set out to illustrate. Hence they lack targets.[18]

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 at its target is wholly a matter of how the world is (at the target), and is

---

[16]Of course, cameras are artifacts, so their function itself is the product of human intention. But the relevant intentions in this case are directed (*de re*) to the picture making device, and not (*de re*) to each picture it produces. By contrast, the artist of a drawing or painting has representational intentions about the individual work she is producing. Still, as Megan Hyska pointed out to me, cameras themselves can be used for different functions— e.g. as devices of duplication (as with a Xerox) or as devices of photography; selecting one such function might indeed turn on the *de re* intentions of the device's controller.

[17]Kjørup (1978) articulates a cognate norm of "seriousness" for pictorial speech acts, though he does not allow that an intentionally lying artist could be using a picture in the same assertoric sense as a sincere one.

[18]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, typically, *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.

independent of artistic skill.

Thus when an artist sets out to create an (assertoric) drawing, she comes to the table with at least two sets of intentions. One is an expressive intention; this is the intention to create a picture which expresses a particular content, regardless of its accuracy. The other is an "aiming" intention; this is the intention to create a picture whose content is accurate at a particular scene, whatever that content happens to be. Assertoric drawing always involves both kinds of intentions, and both play a role in the psychological explanation of a given picture's production.

This concludes my overview of the Three-Part Model. In the next section, I'll argue from cases that some version of the Three-Part Model must be correct.

2   Singular Content vs. Target

At the outset, I characterized the depiction literature as positing singular and attributive content, to which the Three-Part Model added target. A more accurate 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 singular element, when combined with attributive content, determines an accuracy value (like target). For these reasons, a more careful statement of the primary contribution of the Three-Part Model is that, unlike previous accounts, it recognizes both singular content *and* target and distinguishes between them. This position can be characterized by a pair of theses, each of which is defended separately below:

Thesis 1. Assertoric pictures express singular content, which is independent of their attributive content or target (if any).

Thesis 2. Assertoric pictures' contents are (in)accurate only relative to a target scene, which is independent of their attributive content or singular content (if any).

Here, saying that one representational relatum is independent of the other two means that it cannot be *derived* or *determined* from the other two, not that it is unconstrained by them. The point for Thesis 1 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 2, 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. This fact will play a central role in the argument for the Three-Part Model developed below.

In this section I'll argue for the Three-Part Model's two distinctive theses 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 features of these cases.

2.1 Thesis 1: Singular Content

I first argue for for the Three-Part Model's Thesis 1, that pictures have singular contents, independent of their attributive contents or targets (if any).

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 noted, 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: pictures 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 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 *targets* via their *visible parts*. 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:

> P is of O iff (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. Bu otherwise, if it is still within the picture frame, it will count as visible. [19]

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 (as the Three-Partist would have it) 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.

---

[19] The concept of the geometrical projection at work here will vary by system of depiction; some systems rely on linear perspective projection, others isometric projection, and so on (Greenberg 2013; Giardino and Greenberg 2015).

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, pp. 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:[20]

> P is of O iff (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 because the object in question need not be *visible* in the target scene, so long as it is appropriately picked out by the attributive content.[21]

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. My approach is to identify 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 1.

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 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

---

[20]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.

[21]In addition to the kinds of objections pursued below, 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, p. 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, p. 30) and Abell (2009, p. 212).

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, false background beliefs, or simple fantasy 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 see for the first time a cube, named Cubey, 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 that point I 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:
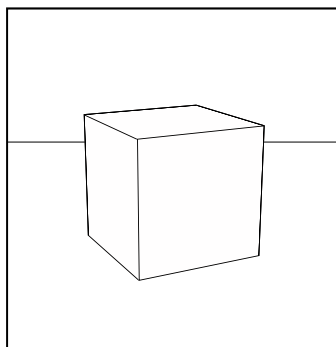
Picture A

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 isn't accurate 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— 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. But now we can see that the synecdochic strategy is directly counter-exampled by the Object Hallucination case. For 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, p. 97) has noted, this cannot work.[23] Applied to the case above,

---

[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.

[23]Lopes (1996, p. 94) loosely characterize the descriptivist view as holding that the a picture is *of* an object when its

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 synechdocic 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.

In sum, there is no valid way to use the attributive content, even together with a picture's target scene, to specify the objects the picture is of. If the mode of specification is too relaxed, allowing for objects beyond those visible in the target, then the picture is predicted to be of objects which in fact it is not. If the specification of objects is limited to those visible in the target, then the picture is predicted not to be of objects which in fact it is.

Lopes (1996, pp. 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. The only feasible response is to hold that for a picture 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 scene that fit the attributive content *to some degree*. There seems to be no chance of picking out Cubey, not only from other cubes, but also from other objects that are sufficiently similar.

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.[25]

---

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.

[25]One might take the arguments here only to support the conclusion that pictures have referents, independent of their attributive contents and targets. But why think that these referents are also part of content? Two reasons prevail.

First, facts about what objects pictures are *of* and 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*. In so far as these facts are reflections of content, as I have assumed, they indicate that content contains and integrates both singular and attributive content.

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

2.2 Thesis 2: Target

I turn now to the second key claim of the Three-Part Model: that pictures are evaluated for accuracy relative to target scenes, which are independent of a picture's attributive content and singular content (if any).

Partial support for this claim comes from the fact that it is possible to evaluate pictures for accuracy relative to *arbitrary* scenes, simply by fiat. Thus, given an accurate picture, I may ask if what it depicts (i.e. its content) *would* be accurate relative to any other possible scene. But if such evaluations are possible, then accuracy is relative to a scene in a manner unconstrained by pictorial content. And so absolute ascriptions of accuracy must implicitly recognize a distinguished such scene— this, I suggest, is the target. Still, such considerations are not decisive with respect to Thesis 2; it remains an open question whether the *target* scene (as opposed to an arbitrary scene of evaluation) is really independent of pictorial content in this way.

Skepticism about the existence of independent targets may take different forms. One approach holds that pictures are evaluated for accuracy relative to target scenes, but these scenes are derived from facts about singular content and attributive content. In essence, on this view, pictorial content *fixes* a picture's target; for instance, the target might be defined as any one of the scenes which satisfies a picture's content. But this hypothesis proves unworkable. The problem is that, if content fixes target, it is hard to see how a picture's content could ever be inaccurate, since accuracy is defined in terms of the match between content and target.

The strategy fairs no better when restricted to the claim that target is determined by singular content alone. For any given collection of singular elements can be recombined to be compatible with a variety of visibly distinct targets. But then pictorial contents which put, as it were, the right objects in the wrong arrangements, will still be counted as accurate even when they are not.

A more plausible and challenging view holds 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 do not themselves instantiate their properties and relations absolutely, but only 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 object "tropes". The idea is that by rooting the singular content of a picture to a particular time

---

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* be accurate at the scene in which Yebuc is visible. This is simply because the content contains the object Cubey, so Cubey helps determine the picture's accuracy conditions. (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.)

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 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. At w and t, C completely overlaps Cubey and has all the same intrinsic properties as Cubey 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:

> The content of P is accurate iff for every world- and time-bound object O in P's singular content, O instantiates the properties associated with it by P's attributive content.

A picture is accurate when all the object instances in its singular content instantiate the properties attributed to them. 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.[26]

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:

> P is of O iff 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 inex-based accounts). At the very least, for a object instance to be part of a picture's singular content, there must be some causal connection between the instance and the creation of the picture; and in many cases this causal connection must be mediated by the representational intentions of the artist.

---

[26]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. The index-free theorist might claim that viewpoints themselves are part of pictures' singular contents. Or they might existentially quantify over viewpoints in the definition of accuracy.

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 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 very close 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 see for the first time a cube, named Cubey, 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 that point I 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. I proceed to draw the situation I take myself to be seeing, producing Picture B.
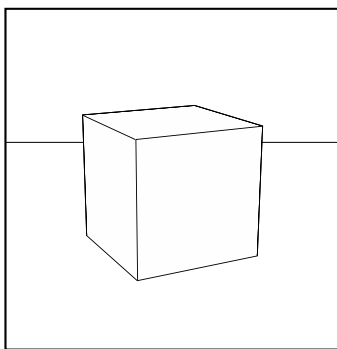
Picture B

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 the content of Picture B would be *accurate*, since 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.[27]

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. As advertised in Section 1, this implies both that two pictures with the same content can have different targets. We can extend the Scene Hallucination case to make these points 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 one of 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

---

[27]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 ⟨*object, scene*⟩ 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 semantic 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 guided by methodological considerations such as ease of use and elegance. My own formalization has the virtue of making the structure of pictorial representation commensurate with that of language.

object and target dissociate from one another, as in the two cases above. Theories which collapse or reduce these two elements lack the resources to account for such variation.[28]

3   Non-Factual Targets

Thus far I have articulated the Three-Part Model and argued for its core tenets. In this section I turn to *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 target. Such examples are prima facie challenges to the Three-Part Model's claim that all assertoric pictures have particular 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; it if it is accurate at a target, that target must be counterfactual. 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

---

[28]Here I should note that I have, for the sake of expository clarity, overstated my case somewhat. The arguments presented establish that singular content is independent of the target scene relative to which a *picture* is assigned an accuracy value in context. I have assumed that it follows that the target scene is the index relative to which a picture's *content* is assigned its accuracy value. This corresponds to my preferred presentation of the Three-Part Model. But an alternative view, compatible with the arguments presented in this section, holds that while the *world* of the target makes up the index of evaluation, the *viewpoint and time* of the target is instead a feature of context relative to which the picture *expresses* content. On this view, the target viewpoint still affects accuracy value, but through a different means. A central motivation for choosing my preferred account over the alternative, is indicated in Section 1: intuitively, we can assess a given picture's content at different viewpoints and worlds— but this would be impossible on the alternative view. A fuller treatment of this issue lies beyond the scope of this paper.

[29]It is important to distinguish architectural *plans* from the kind of drawing here, which (for certain style 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.

already made allowances for such descriptive selection of targets— afterall, this is what happens when I aim at an (actual) scene which is merely relayed to me via description. In the case of counterfactual depiction, the artists 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*, such as that of Stalnaker (1968), Lewis (1973), or 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' intentions pick out a set of possible worlds as their targets in counterfactual depiction.

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. 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 *target set* of scenes in terms of the former notion:

> The content of P is accurate at a target set of scenes S iff
> there is at least one scene s in S such that the content of P is accurate at s.

Recall that when a picture is accurate at its target, we may say that it is *absolutely* accurate, or accurate without relativization. Thus, when a picture's target is non-factual, it is absolutely accurate when there were is some scene in its target such that it is accurate *relative* to 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, but it may very well be accurate at *some* such scene. Our intuitions of accuracy seem to track the latter condition.[30]

The case of GENERIC pictures presents a related and in some ways more demanding challenge to the Three-Part Model.[31] 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 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.[32] At the very least, the idea that we might be

---

[30]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 fiction may not be real, 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 be mean evaluating the picture for accuracy at a set of scenes compatible with the possible worlds specified by that fiction.

[31]Thanks to an anonymous reviewer for pressing the challenge posed by generic pictures, and to Susanna Siegel for discussion of possible solutions.

[32]We must also accommodate the fact that pictorial content is viewpoint relative. Viewpoint relative content can only be assessed for accuracy relative to a viewpoint, so somehow viewpoint must play a role in the definition of accuracy, even for a generic picture. This does not establish that viewpoint must be part of the *target* relative to which a generic picture is evaluated— but the only alternative is that viewpoint is quantified over in the definition of generic accuracy. Indeed, it is natural to think that, for a generic picture to be accurate, its content must accurate at *some* viewpoint located within a "generic" world. The only question is whether the quantification must be built-in as a special feature of the definition of accuracy for generics. In the alternative analysis offered below, I show how such quantification falls out naturally from the framework of non-factual targets, without modifying the definition of accuracy.

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. In this case, the targets of generic pictures are picked out in a distinctive way, by reference to part of an 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 content. Consequently, the content of a generic picture of Obama would be accurate absolutely because it is accurate at some particular possible scene where Obama appears as he normally would.[33]

Note that, on this account, no particular viewpoint, or set of viewpoints need by 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] So some constraints on viewpoint may need to be introduced as well.)

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. For, on one hand, generic pictures aim at targets that are not (necessarily) culled from the actual world, but are related to the actual in reliable ways. And on the other, the targets of generic pictures are *sets* of scenes; a generic picture is true at such a set when its content is accurate at some member of the set.

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— everything from photography and drawing from life to counterfactual and generic depiction. It is clear that much remains to be said, but I hope the sketches offered here suggest that the resources of the Three-Part Model are well-suited

---

[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 res triction is necessary here as well.) At any rate, this a first-pass analysis of this kind of depiction.

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

to the task.

4 Conclusion

In this paper 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 part of the standard account of the semantics of sentences within philosophy of language and linguistics (Kripke 1972; Lewis 1980; Kaplan 1989). These bear clear analogies to the notions of singular content, attributive content, and target in the Three-Part Model.

Similar comparisons may also apply to the cases of visual perception and mental imagery. In the philosophy of perception, it is commonly (though not universally) thought that perceptual states have both singular and attributive content (Burge 2010; Siegel 2011). Following the spirit of Cummins (1996), I suspect the notion of target also finds a home here. Its plausible that much of the argument of this paper— stemming from the Scene Hallucination case— may be carried over to the domain of perception with little loss. If it can, then a distinction between content and target would be necessary for perception as well. Further, the mere fact that mental images, like drawings, can represent both past, present, actual and counterfactual scenarios, strongly suggests the work of a counterpart notion of target for imagery. Meanwhile, perception, like photography, seems always to target the scene before the eye. Thus broad analogies in the interplay of content and target emerge between mental visual representation and depiction.

If these speculations are on track, then, across a range of cases, from pictures, to perception, to language, we find the find the same basic elements that make-up the Three-Part Model: a representation expresses both (i) singular content, and (ii) attributive or predicative content; and (iii) this content is evaluated for accuracy (or truth) at a target index. Though the rules that associate representations with contents will vary dramatically from from case to case, I conjecture that the same basic semantic architecture is characteristic of all sufficiently complex representational systems.

References

Abell, C. (2009). "Canny Resemblance". In: *The Philosophical Review* 118.2, pp. 183–223.

Blumson, B. (2009). "Pictures, Perspective and Possibility". In: *Philosophical Studies* 149.2, pp. 135–151.

Burge, T. (1991). "Vision and intentional content". In: *John Searle and his Critics*. Ed. by E. Lepore and R. van Gulick. Blackwell.

— (2010). *Origins of Objectivity*. Oxford: Oxford University Press.

Burge, T. (2014). "Reply to Rescorla and Peacocke: Perceptual Content in Light of Perceptual Constancies and Biological Constraints". In: *Philosophy and Phenomenological Research* 88.2, pp. 485–501.

Camp, E. (2007). "Thinking with Maps". In: *Philosophical Perspectives* 21.1, pp. 145–182.

Cummins, R. (1996). *Representations, Targets, and Attitudes*. Cambridge, MA: MIT Press.

Frege, G. (1879). *Begriffsschrift, ein der arithmetischen nachgebildete Formelsprache des reinen Denkens*. Translated as *Concept Script, a formal language of pure thought modelled upon that of arithmetic*, by S. Bauer-Mengelberg in J. vanHeijenoort (ed.), *From Frege to Gödel: A Source Book in Mathematical Logic, 1879–1931*, Cambridge, MA: Harvard University Press, 1967. Halle/Saale: Verlag L. Nebert.

Geach, P. (1965). "Assertion". In: *Philosophical Review* 74, pp. 449–65.

Giardino, V. and G. Greenberg (2015). "Varieties of Iconicity". In: *Review of Philosophy and Psychology* 6.1, pp. 1–25.

Goodman, N. (1968). *Languages of Art: An Approach to a Theory of Symbols*. Indianapolis: Bobbs-Merrill.

Greenberg, G. (2013). "Beyond Resemblance". In: *Philosophical Review* 122.2, pp. 215–287.

— (2016). "The Geometry of Pictorial Representation". Manuscript.

Hagen, M.A. (1986). *Varieties of Realism: Geometries of Representational Art*. Cambridge: Cambridge University Press.

Hopkins, R. (1998). *Picture, Image and Experience: A Philosophical Inquiry*. Cambridge University Press.

Howell, R. (1974). "The logical structure of pictorial representation". In: *Theoria* 40.2, pp. 76–109.

Huffman, D.A. (1971). "Impossible Objects as Nonsense Sentences". In: *Machine intelligence*, p. 295.

Hyman, J. (2012). "Depiction". In: *Royal Institute of Philosophy Supplement* 71, pp. 129–150.

Kaplan, D. (1968). "Quantifying In". In: *Synthese* 19.1-2, pp. 178–214.

— (1989). "Demonstratives: An Essay on the Semantics, Logic, Metaphysics, and Epistemology of Demonstratives and Other Indexicals". In: *Themes from Kaplan*. Ed. by Joseph Almog, John Perry, and Howard Wettstein. Oxford University Press, pp. 481–563.

Kjørup, S. (1978). "Pictorial speech acts". In: *Erkenntnis* 12.1, pp. 55–71.

Koenderink, J.J. and A.J. van Doorn (2003). "Pictorial space". In: ed. by H. Hecht, R. Schwartz, and M. Atherton. MIT Press.

Korsmeyer, C. (1985). "Pictorial assertion". In: *The Journal of aesthetics and art criticism* 43.3, pp. 257–265.

Kratzer, A. (2012). *Modals and conditionals: New and revised perspectives*. Vol. 36. Oxford University Press.

Kripke, S. (1972). "Naming and Necessity". In: *Semantics of Natural Language*. Ed. by Donald Davidson and Gilbert Harman. Revised edition published in 1980 as *Naming and Necessity* (Cambridge, MA: Harvard University Press). Dordrecht: Reidel, pp. 253–355, 763–769.

— (1977). "Speaker's Reference and Semantic Reference". In: *Midwest studies in philosophy* 2.1, pp. 255–276.

Lewis, D. (1973). *Counterfactuals*. Blackwell Publishers.

— (1978). "Truth in Fiction". In: *American Philosophical Quarterly* 15, pp. 37–46.

— (1980). "Index, Context, and Content". In: *Philosophy and Grammar*. Ed. by Stig Kanger and Sven Ohman. Dordrecht: D. Reidel, pp. 79–100.

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

McDowell, John (1984). "De re senses". In: *The Philosophical Quarterly (1950-)* 34.136, pp. 283–294.

Nickel, Bernhard (2016). *Between logic and the world: An integrated theory of generics*. Oxford University Press.

Peacocke, C. (1992). *A Study of Concepts*. The MIT Press.

Rescorla, Michael (2008). "Predication and cartographic representation". In: *Synthese*.

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: *Ifs. Conditionals, Belief, Decision, Chance, and Time*. Dordrecht, D. Reidel Publishing Company, pp. 41–55.

Willats, J. (1997). *Art and Representation: New Principles in the Analysis of Pictures*. Princeton: Princeton University Press.