Review of Philosophy and Psychology
https://doi.org/10.1007/s13164-020-00513-6

Semantics of Pictorial Space

Gabriel Greenberg[1]

Accepted: 30 September 2020Published online: 02 January 2021

Abstract

A semantics of pictorial representation should provide an account of how pictorial signs are associated with the contents they express. Unlike the familiar semantics of spoken languages, this problem has a distinctively spatial cast for depiction. Pictures themselves are two-dimensional artifacts, and their contents take the form of *pictorial spaces*, perspectival arrangements of objects and properties in three dimensions. A basic challenge is to explain how pictures are associated with the particular pictorial spaces they express. Inspiration here comes from recent proposals that analyze depiction in terms of geometrical projection. In this essay, I will argue that, for a central class of pictures, the projection-based theory of depiction provides the best explanation for how pictures express pictorial spaces, while rival perceptual and resemblance theories fall short. Since the composition of pictorial space is itself the basis for all other aspects of pictorial content, the proposal provides a natural foundation for further pictorial semantics.

A semantics of pictorial representation must provide an account of how pictorial signs are associated with the contents they express. Unlike the familiar semantics of spoken languages, this problem has a distinctively spatial cast for depiction. Pictures themselves are two-dimensional artifacts and their contents take the form of **pictorial spaces**: perspectival, three-dimensional arrangements of objects and properties. A basic challenge for pictorial semantics is to explain how pictures express the spaces which they do. In this essay, I argue that, for a central class of pictures, the expression of pictorial space is founded upon a relation of geometrical projection between a picture and the space it expresses as content. The result is an approach to pictorial representation that I'll call **projection semantics**.

Artists and scholars have long known that projection is an ideal recipe for producing pictures. More recently, several authors have proposed that there are also constitutive connections between projection and what a picture depicts (Hyman 2006,

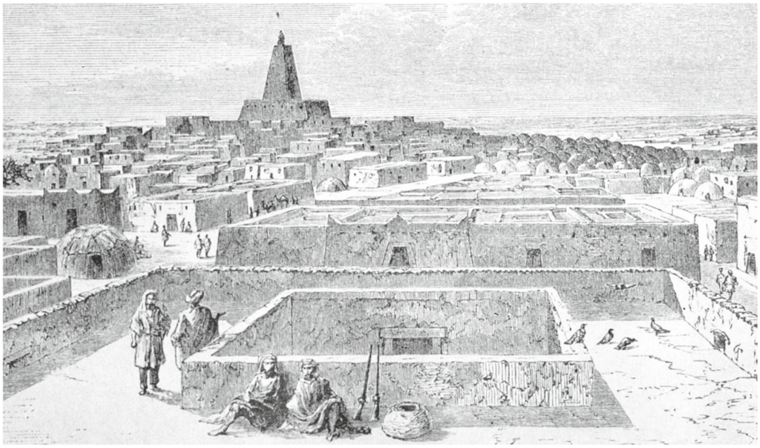
✉ Gabriel Greenberg
gabriel.greenberg@gmail.com

1 Department of Philosophy, University of California, Los Angeles, CA, USA

ch. 5; Hyman 2012, §5; Kulvicki 2006, ch. 3; Greenberg 2011, chs. 1-3, 2013). As I'll show here, this depiction-projection link can be recast as as an explicit semantic principle which associates pictures with constraints on their accuracy conditions. My aim in this essay is to turn such a semantic principle to the problem of pictorial space. I will argue that a principle of this kind provides the best explanation for how pictures in a wide range of styles express pictorial spaces, in particular the layout of basic features in viewpoint-centered space. Since the composition of pictorial space is itself the basis for all other aspects of pictorial content, the projection principle offers a foundation for pictorial semantics.

The argument is developed in stages. First, in Sections 1-3, I'll show how projection semantics provides an explanatory foundation for the expression of pictorial space, using depiction in linear perspective as my prime example. This result is generalized in Section 4 for a wide range of systems of depiction beyond linear perspective, which include the central class of "realistic" or "mechanical" pictorial styles. Then in Section 5, I'll make the case that facts about the human visual system help explain why humans would favor a pictorial signaling system based on projection of the kind I've proposed. In Section 6, I'll argue that alternative resemblance and perceptual theories of depiction, when considered as accounts of the organization of pictorial space, fall short as compared to a projection semantics. Section 7 examines the prospects for projection semantics with respect to a range of more organic representational styles, such as caricature and stylization. Section 8 surveys the essay's conclusions.

1 Pictorial Space



Picture A

In 1860 a visitor to the city of Timbuktu, Mali, published this engraving, which I'll call Picture 1.[1] It depicts a collection of buildings as laid out in a loose grid; it

[1]The artist was probably Dieudonnée Lancelot. See: https://commons.wikimedia.org/wiki/File:Barthtimbuktu.jpg.

depicts two people in the foreground as sitting down; and it depicts what is likely the Djinguereber Mosque in the background as having a certain angular shape. What objects and individuals the picture depicts, and what properties and relations it depicts them as having, are all reflections of the picture's **content**. The content of a picture corresponds to what's happening in it, or the situation it represents, which in turn determines the set of conditions under which the picture is accurate. Officially, I'll say that pictures **express** their contents, or informally, that they "depict" or "represent" elements of their contents.

By **pictures** here I mean a broad class of public sign whose representational properties are in a rough sense distinctively visual. They are characterized in part by their two-dimensional spatial structure, and in part by their content, which, as I'll argue, takes the form of a three-dimensional space anchored at a viewpoint. Canonical examples of pictures include architectural and engineering drawings, figurative paintings, functional illustrations sketches and illustrations, photographs, as well as many kinds of maps.[2]

While the category of pictures include works of visual *art*, it is not limited to them.

The task for a **semantics** of pictures is to determine the rules, if any, by which pictures, in context, may be associated with the contents they express.[3] In engaging this task, philosophical theories of depiction have tended to focus on how and why pictures depict one object rather than another, or none at all, or why they attribute high-level properties like *being a plant* or *sitting down* to these objects. Thus they tend towards providing conditions for the ascriptions of content of the following form, where P is a picture, and x and F are an object and (high-level) property respectively:[4]

$$P \text{ depicts } x \text{ as } F$$

In this section, I wish to argue for a reorientation of pictorial semantics away from the traditional focus on "depiction-as" ascriptions, and towards a more holistic view of pictorial content as comprising a kind of space.

The problem is that any true "depiction-as" ascription, while undoubtedly a reflection of pictorial content, cannot provide a complete description of a picture's content. Pictorial content invariably involves many objects, not just one x, and many properties attributed, not just one F. Thus Picture A not only depicts two people in the foreground as sitting down; it also depicts the pigeons as perched on the roof, the mosque as oriented at a certain angle in the background, and so on for the many other

[2]A caveat for the case of photography: because the process by which photographs are produced is to such a high degree mechanical, reasonable questions have been raised about whether photographs should count as representations at all (Scruton 1981). Here I will proceed on the assumption that photographs *are* representations: they are pictures, they express contents, and their contents encode conditions of accuracy. But the inclusion of photographs is not central to my theoretical aims.

[3]This use of "semantics" doesn't distinguish pictorial semantics from pragmatics, instead referring to the general interpretive framework for pictures, whatever it may be.

[4]Prominent analyses of pictorial representation which focus on, or culminate in conditions on "depiction-as" include Abell 2009, 217; Goodman 1968, 26-31; Hopkins 1998 76-77; Lopes 1996, 151-153; and Walton 1973, 312-15. Meanwhile, literature on the psychology of pictorial perception has followed the holistic concerns with pictorial space advocated here. See footnote 5.

objects and properties depicted. In this way, the content of a picture seems to determine a large, possibly infinite, set of such "depiction-as" ascriptions. But it would be unilluminating simply to identify pictorial content with a list of ascriptions like this. Rather, it seems, we should think of the content of a picture as the densely packed situation which in turn gives rise to such ascriptions— the situation relative to which each such ascription is true.

Whats more, the myriad "depiction-as" ascriptions licensed by a picture's content posses a kind of spatial unity. For all the objects depicted by Picture A (or, all objects that figures in a true "depiction-as" ascription about Picture A), Picture A also depicts those objects in spatial relations to one another. It depicts the pigeons as *on top of* the roof, the people sitting down as *to the left of* the pigeons, the Djinguereber Mosque as *further away than* the people sitting down, and so on. Thus the properties attributed by a picture's content appear to form a kind of integrated spatial lattice. It now seems that the "situations" underlying pictorial content should be understood as cohesive spatial distributions of objects and properties. Such considerations have led some to conceive of the contents of pictures as **pictorial spaces**: three-dimensional spatial arrays populated with individuals, properties, and relations (e.g. White 1957; Howell 1974; Wollheim 1987; Rogers and et al 2003).[5]

Pictorial space plays a foundational role in pictorial semantics, because all pictorial content, including high-level property ascriptions, are anchored in pictorial space. For each high-level property F that a picture's content attributes to some object, there is always a spatial implementation of F that the picture also attributes. For example, Picture A doesn't merely depict the two people in the foreground as sitting down, it depicts them as sitting down in a particular posture, at a particular angle and relative distance from the viewpoint. This is true for every other object and property represented and it holds with a determinate degree of precision for any given case. Generalizing, it seems to be impossible to represent something in a picture without also attributing to it such determinate spatial locations and features. These facts distinguish pictorial representation from language, where, for example, the sentence "two people are sitting down" does not imply any particular posture or orientation of the sitters (cf. Dretske 1981, ch. 6; Hopkins 1998, 26-7). Indeed, the essentially spatial character of pictorial content seems to be a central mark of pictorial representation.

Of course, pictures are often used to convey rather narrow information, far short of a full pictorial space, which might indeed be captured in a simple "depiction-as" ascription. If you ask me what I had for breakfast, and I send you a detailed drawing, what I am conveying by the picture is arguably simple: I had toast with butter and jam. It is in this spirit that Korsmeyer (1985) analyzes what a picture "asserts" in terms of simple subject-predicate ascriptions. But we should not confuse what a picture, on a given occasion, is used to *convey* (or assert) with the content it more

[5]Note that the term "pictorial space" has often been employed by psychologists to refer to the space that one *perceives* when looking at a picture (e.g. Rogers 1995; Ittelson 1996; Koenderink and van Doorn 2003; Cutting 2003; DeLoache et al. 2003; Vishwanath et al. 2005). Here I use it to describe the form of a picture's content, independent, at least in principle, from the perceptual content that the picture elicits from viewers.

generally *expresses.* The picture of my breakfast expresses more than the presence of toast. It also depicts the type and shape of the toast, the distribution of bread crumbs, the lighting of the background, and so on. The picture's content includes the entire pictorial space populated with all of these objects and their properties, and more. This is why I could have used the same picture to convey any of these more specific ascriptions, but not arbitrary ascriptions divorced from elements within the pictorial space. And it is this spatially unified content that determines the conditions under which the picture itself is accurate, quite apart from the specific message it may have been used to communicate.

If pictorial spaces are the contents of pictures, then pictorial spaces must determine accuracy conditions. To this end, let us say that a pictorial space is **realized** by a particular concrete scene to the extent that the composition of the scene reflects the structure and constituents of the space. I'll define a **scene** formally as a pair containing (i) a possible world and (ii) a viewpoint located in that world— what I'll call a **viewpoint-centered world**. Then we may say that a pictorial space is accurate at a scene if and only if the scene is a realization of that space.[6] Note that a given pictorial space can be realized by indefinitely many scenes. This is because pictorial spaces are always partly indeterminate— if nothing else, they are silent with respect to occluded objects, and objects outside the window of the picture plane— while scenes, as I define them, are fully determinate in all possible ways.

I turn next to the central concern of this essay, the structure of pictorial space. As many authors have noted, pictorial spaces are in some clear sense *perspectival* (see e.g. Budd 1996, 169-70; Hopkins 1998, 27; Casati and Giardino 2013). That is to say, the objects and properties which inhabit pictorial space are all located relative to a central perspective or point of view. The viewpoint-relativity of pictorial space is exhibited in a variety of ways. For example, it arises in relations of *depth*: Picture A depicts the Djinguereber Mosque as being *further from* the viewpoint than the men in the foreground. And importantly for this essay, it is reflected in properties of *direction*: Picture A depicts the minaret of the Djinguereber Mosque as directionally *above* the viewpoint, the pigeons on the roof as to the *right* of the viewpoint, and so on. Critically, such directional relations are not limited to crude cardinal directions, but appear to locate every part of every depicted surface in a quantitatively distinct direction from the viewpoint.

These observations help make clear why most maps should also count as falling within the natural class of pictorial representations. Setting aside the linguistic and symbolic tags with which they are typically adorned (Greenberg 2019), maps, like other pictures, exploit flat, marked surfaces to represent 3-dimensional spaces. While the pictorial spaces expressed by maps tend to be relatively indeterminate with respect to depth (but not flat), they locate their subjects in a clear set of quantitatively fine-grained directions. And crucially, these spaces inhabit viewpoints, displaying their geographic subject matters from above, and not from the side or below, or any oblique angle. Such viewpoints tend to be more planar than point-like, a distinction I return to in Section 4, but are nonetheless vividly perspectival.

[6] See Peacocke (1992, 62) for analogous remarks about perceptual space.

The informal notion of viewpoint at work here can be regimented with a geometrical one. Henceforth, by **viewpoint**, I mean a particular oriented location in space and time, in a particular possible world. The concept of an "oriented location" is definable within the framework of projective geometry, a matter I turn to in the next section. A viewpoint need not be tied to any real, implied, or even imagined viewer. It is simply an abstract, spatio-temporal index, what Hyman (2006) has called an "objective eye."

Whereas viewpoints are token oriented locations, it is also useful to identify **general viewpoints** as *types* of oriented locations. This is the kind of viewpoint which is implicit in a picture's content, and can be perceptually inferred by looking at a picture, even without knowledge of the token viewpoint it is intended to be evaluated at. All objects in a pictorial space are given their location in relation to such a general viewpoint, but the general viewpoint itself is not tied to any particular location at a particular possible world. Instead, a given general viewpoint may be instantiated by any number of (token) viewpoints, which do have definite locations. A pictorial space is realized by a scene just in case, when the general viewpoint of the space is aligned with the token viewpoint of the scene, the positions of objects and properties specified by the pictorial space thereby correspond to the actual locations of those objects and properties within the scene.

To capture the relation between viewpoints and the objects a picture depicts, I adopt a model inspired by Peacocke's (1992, ch. 3) analysis of perceptual content. A general viewpoint, or oriented location type, serves as the origin or "center" of a pictorial space. Then each object and property in the space is assigned a location relative to this center. Locations, in turn are specified by a direction, and typically a depth from the viewpoint, analogous to the locations of a polar coordinate system. As a whole, pictorial space fills a three-dimensional region with object and properties, whose locations are given by a direction and depth from a general viewpoint.[7]

The spatial conception of pictorial content is compatible with models of content based on the possible world framework. The content of a picture can be thought of as a set of scenes, understood as world-viewpoint pairs (Ross 1997, ch. 5; Blumson 2009; Greenberg 2011; Abusch 2015). Collectively these pairs define all the possible conditions and viewpoints relative to which the picture is accurate. A pictorial space corresponds to the sets of scenes for which the viewpoint realizes the general viewpoint of the space, and the world realizes the constituents of the space. The scenes which belong to the content of a picture, on this model, will all coincide precisely insofar as the picture's represented space is determinate, and diverge insofar as it is indeterminate.

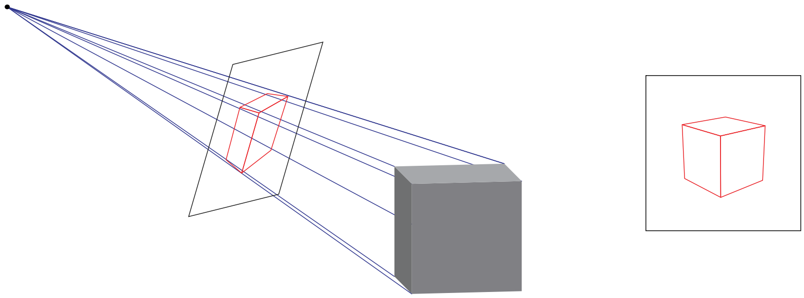
Thus far I have defended and refined the assumption that pictorial contents take the form of viewpoint-centered spaces. In doing so, I've suggested an approach to pictorial semantics that prioritizes the expression of pictorial space over, or at least along side, the traditional attribution-based concerns of the theory of depiction. Of course, what is distinctive about pictorial representation is not just the content pictures express, but the *way* they express it. In particular, there seems to be a precise

[7] This definition is informal, but will suffice for present purposes. See Greenberg (2020) for a more in-depth discussion.

and direct correspondence between the geometrical organization of lines and colors on the two-dimensional picture plane, and the three-dimensional arrangement of features in pictorial space. What I will call the **problem of pictorial space** is the problem of explaining how pictures are systematically associated with pictorial spaces. Resolving this explanatory demand is a base-line for any pictorial semantics. In what follows, I describe the foundations of such an account.

2 Projection Semantics

At least as early as the Italian Renaissance, scholars recognized important connections between depiction on one hand, and geometrical projection on the other. **Geometrical projection** describes a family of methods for transposing a three-dimensional scene onto a two-dimensional surface via a system of spatial relations established relative to a viewpoint. The result, as illustrated in the familiar perspective projection below, is a picture: a determinate arrangement of line and color in two-dimensions.



Projection is traditionally conceived as an idealized method for the construction of pictures. Such methods may be realized by various practical recipes for artists to carry out. Projection is an *idealized* method not only because it is a geometrical, rather than physical, transformation, but also because not all pictures conform with projection. Drawings that are poorly executed or otherwise inaccurate tend to depart from the projective norm, but are nonetheless pictorial representations. Instead, the connection in question seems to relate projection to *accurate* depiction. Thus, creating a picture in accordance with geometrical projection seems, all things equal, to be a recipe for creating an accurate picture; conversely, if a picture is not a projection of a scene, then it seems it cannot be an accurate picture of that scene, at least for a core class of "mechanical" and "realistic" drawing styles.

But accuracy, like truth, is a semantic concept, and this suggests that projection itself has semantic import for pictures. Inspired by observations like these a number of scholars have proposed to reverse the traditional conception of projection and depiction (Hyman 2006, ch. 5; Hyman 2012, §5; Kulvicki 2006, ch. 3; Greenberg 2011, chs. 1-3; Greenberg 2013; Abusch 2015, §2, §5).[8]

[8] Another group of authors, including (Peacocke 1987; Budd 1996), and (Hopkins 1998), enlist projection within the framework of the resemblance theory of projection; I address these theories in Section 6.

Instead of thinking of projection as a norm for the *construction* of pictures, they conceive of projection as a norm for *interpreting* pictures. For Hyman (2006, ch. 5) and Greenberg (2013), this idea is expressed as a projective requirement on the relation of depiction or accurate depiction. Kulvicki (2006, ch. 30) uses projection define a level of pictorial content which exhibits certain geometrical properties.

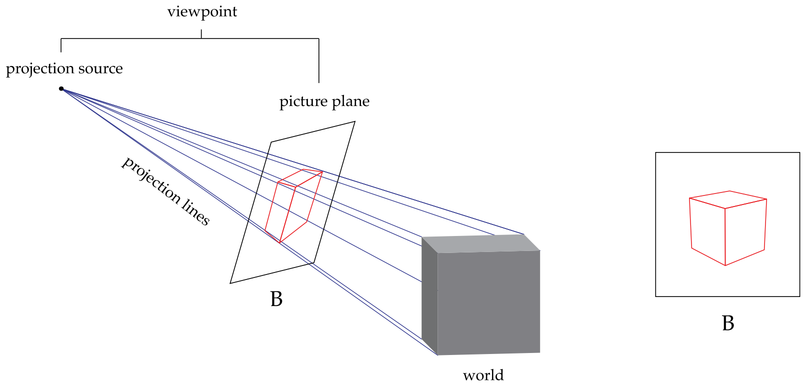
My plan in this essay is to enlist the core idea of a projective constraint on depiction to address the central problem of pictorial space. This new explanatory work will require a reformulation of the underlying idea in explicitly semantical terms, encapsulated in what I will call the **Projection Principle**: for a picture to express a content, the picture must be a projection of that content. The result is an approach to pictorial representation I will call **projection semantics**. In developing this proposal, I continue to focus on the core class of "mechanical" or "realistic" systems of depiction, setting aside for now the more organic systems of stylization, caricature, and their ilk; I return to these in Section 7.

According to the Projection Principle, a picture can only express a given space as content if the picture itself is a projection of that space; the type of projection will vary for different kinds of picture. To interpret a picture is, in part, to determine the type of space that it must be a projection of. Projection here is understood as a purely geometrical relation between a possible scene and a picture plane, not a physical process for producing a token picture. The Projection Principle holds that a picture's content is the space it *purports* to be a projection of, not the situation which actually caused it.[9] Since this approach defines pictorial content in purely geometric terms, without reference to a perceiver, it makes the content of a picture independent, at least in principle, from the perceptual content which the picture may elicit in a real or ideal viewer. In Section 6, I'll argue that pictorial content and the perceptual content provoked by a picture regularly come apart.

Still, such a projection semantics highlights important analogies between pictorial representation and the natural encoding of information in the retinal image. The distribution of light intensity delivered to the retina carries information about the shapes of the objects that reflected it. To extract this information, a central function of the visual system, is to determine what kinds of objects must have caused the retinal image, given the laws of optics and regularities in the environment. On the present view, pictures carry content in a manner analogous to the way the retinal image carries information. To calculate a picture's content is to work how it must have been produced— only, the notion of "production" here is not that of causal production, but of geometrical projection. In this sense, pictures are the abstract traces or footprints of the spatial contents they express (Leyton 1992, ch. 2).

[9]There are other aspects of pictorial content, in particular those that reflect pictorial reference to individuals, which do depend on causal connections between a picture and elements of the context in which it is created.

In order to formulate the Projection Principle more precisely, the concepts of projection and viewpoint must be elaborated.[10] There are many different types of projection, a theme I return to in Section 4. For now, a simple species of *linear perspective with line drawing* is illustrated below. Here we begin with a concrete three-dimensional region of spacetime (possible or actual), which I'll think of as a possible world. In the example below, the world contains only a cube. Next, a **projection source** is located within the space of the world. A projection source is thought of as a geometric point in space and time. This in turn defines a system of **projection lines**, which link each point in the world to the source.[11] Finally, a **picture plane** is introduced into this spray of lines, and they are used to map spatially distributed features of the scene back to surface features of the picture plane itself— in this case, the lines of the line drawing. The result of such a projection is a **marked picture plane**, displayed at right below.



We are now in a position to define a **viewpoint** as a pair of indices, the first of which gives the spatio-temporal location of the projection source, and the second the spatio-temporal location of the picture plane.[12] Together these make up what I have called an "oriented location," where the projection source fixes the location, and the position of the picture plane relative to the projection source determines its orientation.[13] Given a world and a viewpoint of this kind, a method of projection will always deliver a picture. I propose that it is viewpoints in this sense which figure in the formal definition of scenes as viewpoint-centered worlds. Thus methods of projection can be thought of as recipes for deriving pictures from scenes.

[10] The following presentation of the idea of projection draws from Dubery and Willats (1972) and Sedgwick (1980), Willats (1997, ch. 2), and Greenberg (2013) among others.

[11] The projection lines in the diagram below are only a sample of the full array.

[12] Note that the temporal location of the source and plane must be the same. Also, the picture plane itself must define at least up-down and left-right orientations; front and back can typically be defined by the relative position of the projection source.

[13] The notion of *general viewpoint* introduced in the last section can be defined as the pair of an abstract projection source and picture plane, specified only in terms of their positions relative to one another.

All methods of projection are defined by at least two conditions. The geometry of the projection lines is fixed by a **projection condition**. In the case of linear perspective projection, this condition stipulates that the projection lines are straight, and that they converge on a point. But geometry alone is insufficient to define a picture, because something must determine which markings in particular end up on the picture plane. It is the function of the **marking condition** to specify how basic features in the scene are mapped to basic marks on the picture plane (Willats 1997, chs. 4-6).[14] For illustration, a very simple system of line drawing might invoke the following marking condition:

$$\textit{edge} \rightarrow \textit{black}$$
$$\textit{surface} \rightarrow \textit{white}$$

This condition is interpreted to require that if a point lies on an unoccluded edge in a scene (relative to the viewpoint), then it projects to a black point on the picture plane, and if a point lies on an unoccluded surface, then it projects to a white point.[15] Once the arrangement of projection lines is fixed, it is this rule that determines which features of the scene (edges or surfaces) are registered as marks (white or black) by their projective counterparts on the picture plane.

Next, I wish to refine the description of the Projection Principle according to which pictures are projections of their spaces. In order to make this idea precise, it is most natural to define projection as a relation between a picture and a complete *scene*, rather than a pictorial space. So in its official articulation, the Projection Principle holds that, given a picture and the pictorial space it expresses, for every scene which realizes that space, the picture is a projection of that scene. (I'll still talk of a picture being a "projection of its space," meaning that it projects from every scene compatible with that space.) Thus the principle indirectly constrains what space a picture may express by directly constraining the set of scenes which may realize such a space.

A final factor is the existence of different *forms* or *systems* of depiction, a matter I return to in Section 4. The Projection Principle analyzes such systems in terms of different forms of projection. Thus, officially, the Projection Principle claims that, for a picture to express a pictorial space as content *relative to a system*, that space must project back to the picture itself, according to the method of projection *characteristic of that system*.

[14]Willats (1997, 4-20) organizes methods of projection into "projection systems" (≈ my "projection condition") and "denotation systems" (≈ my "marking condition"). See Durand (2002) for an extension and critical discussion of this approach.

[15]Many of the illustrative pictures in this essay also include "blank space", unfilled white areas around the central figure that extend to the border of the page. In principle, such blank space may be thought of as outside the picture plane, not the product of projection, or it may be analyzed as the projection of deep space or void. In the second case, a theory may treat blank space as a different kind of mark than white (though the difference may not be visible). The issue awaits further discussion. I gloss over such complications in the text.

Putting these ideas together, we can now state the Projection Principle in schematic form.[16]

The Projection Principle (for content)
If a picture P expresses pictorial space C, relative to system S,
then for any scene E, centered at viewpoint V, which realizes C:
P is a projection of E from V, via the method of projection determined by S.[17]

The Projection Principle is a *semantic* constraint, because it restricts what content a given picture may express. It narrows the range of scenes compatible with the content of a picture by ruling out exactly those scenes which fail to project to the picture, and allowing in all others. Alternatively, the semantic contribution of the Projection Principle can be thought of as a layer of content itself. The principle requires that a picture's content include exactly those (quite abstract) spatial and chromatic properties had in common by the myriad possible scenes which could project to the same picture. Together, these properties make up what Kulvicki (2006, ch. 3) has called a picture's **skeletal content**. Framed in this way, the Projection Principle can be understood as the requirement that a picture's overall content be compatible with its skeletal content. In the next section, I'll show how skeletal content corresponds to a natural structural constraint on pictorial space.

[16] With some translation, the Projection Principle is comparable to the various projection-based proposals in the literature. (i) *Kulvicki:* Kulvicki posits a basic layer of pictorial content, called **bare bones content**: "whatever scenes could have resulted in a particular [linear perspective picture] via a perspective projection count as parts of the [bare bones] content of the picture" (Kulvicki 2006, 59). He further holds that bare bones content constrains a picture's overall content. The result is equivalent to the Projection Principle for content. However, in order to preserve the property of "transparency," Kulvicki takes a non-standard view of pictorial structure, with the result that Kulvicki's principal is logically weaker than the Projection Principle. Here I believe the Projection Principle captures intuitive distinctions in content that Kulvicki's principle misses; see footnote 33 and the Appendix for a detailed discussion.

(ii) *Hyman:* the Projection Principle is comparable to the *Occlusion Shape Principle* proposed in Hyman 2006, ch. 5. The *occlusion shape* of an object is, roughly, the shape optically projected from an object to an intersecting plane along a given "light of sight." The Principle is then formulated as follows: if a part P of a picture depicts an object O, then "the occlusion shape of O and the shape of P must be identical" (Hyman 2006, 81). Assuming that the sum of objects a picture depicts always determine a specific set of scenes, and reading "depicts an object" as *accurately depicts an object*, the principle expresses a constraint very much like the Projection Principle for accuracy, at least with respect to overall shape. In the end, however, Hyman's principle is tied to an optical interpretation of projection, which significantly narrow its range relative to the Projection Principle; see footnote 32.

(iii) *Greenberg:* for Greenberg (2013), "accurate depiction" of a scene may be understood in the present nomenclature simply as *accuracy at a scene*; Greenberg terms this scene the "referent" of a picture; I understand this simply to be the index relative to which a picture is evaluated. Also see the discussion later in this section on Greenberg's view that projection is sufficient for accuracy.

(iv) *Howell:* a final proposal not mentioned in the text is that of Howell (1974, §3), who defines a picture's *picture space* as a type of spatial array for which the picture could be a projection. However, because Howell's goal is to provide a formal semantics for *sentences* which attribute content to pictures, rather than for pictures themselves, the semantic properties of pictures are never fully set out.

[17] To state the principle formally, let $[\![P]\!]_{S,c}$ denote the content of a picture P, relative to a system S and context c. And let $proj_S(\cdot)$ be a projection function from scenes to pictures, for the method of projection characteristic of S. According to the Projecting Principle, given a system S and context c, for any picture P in S: $\forall w \forall v$: if $\langle w, v\rangle \in [\![P]\!]_{S,c}$ then $proj_S(w, v) = P$. Equivalently: $[\![P]\!]_{S,c} \subseteq \{\langle w, v\rangle \mid proj_S(w, v) = P\}$.

It is useful to characterize the Projection Principle not only as a constraint on pictorial space (or content), as above, but also as condition on pictorial *accuracy*. This brings the principle in line with the familiar rules of linguistic semantics, which characterize content, but typically deliver truth-conditions (or conditions of satisfaction) as well. My starting point is the notion of *relative accuracy*: a picture is **accurate at** a scene, relative to a system and context, when the content it expresses holds at, or is compatible with that scene— much the way linguistic propositions are thought hold (or not) at a world (Greenberg 2018, §1). [18] When contents are conceived as spaces, the holding relation is just the relation of realization between a scene and a space. Note that, for a picture to "accurately depict" a scene is simply for the picture to be accurate at the scene.

Accuracy here is meant to be the pictorial analogue of truth for sentences. A picture is accurate at a scene when, to the extent that its content 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 their 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.

We may then state the Projection Principle in terms of accuracy; assuming that a picture is accurate at all and only the scenes compatible with its content, the two formulations are equivalent:

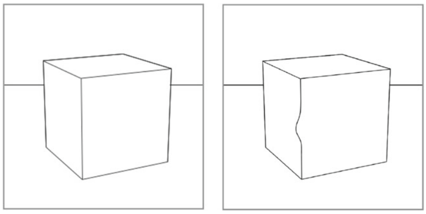
The Projection Principle (for accuracy)
If a picture P is accurate at a scene E, centered at viewpoint V, relative to system S, then P is a projection of E from V, via the method of projection determined by S.[19]

For now, we may think of the semantic significance of the Projection Principle as primarily negative. For any given picture, the principle correctly predicts which contents that picture *could not* express, or of which scenes it *could not* be accurate. For example, given Picture A, apparently depicting a complex of buildings laid out in space, the Projection Principle allows us to rule out innumerable bizarre interpretations. We know from the principle alone that Picture A does not express a space consisting only of an empty room, or a seaside view, or a herd of elk; correspondingly, it would be inaccurate at any scene satisfying these descriptions. Such negative assessments of content may be arbitrarily fine-grained. For example, by the Projection Principle, we know that the geometrically divergent images below necessarily have divergent contents; we also know that Picture D could not accurately depict a scene containing only a (perfectly regular) cube, while Picture C could— with the same result, of course, for even more subtle differences.

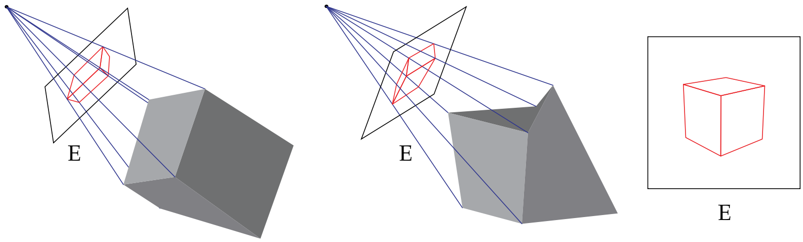
---

[18] 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. Imperatival pictures, for example, do not function to be accurate, but they still have the same kind of spatial content and accuracy conditions as any other picture.

[19] Formally: given a system S and context c, then for any picture P in S, and any world w, viewpoint v: if P is accurate$_{S,c}$ at $\langle w, v\rangle$ then $proj_S(w, v) = P$.



The constraint which the Projection Principle would impose is in an important sense only skeletal, and the principle does not aim to provide a complete theory of depiction. The point is made vivid in the diagram below. Picture E seems to express a space inhabited by a cube, yet the picture itself can be projected from innumerably many other non-cube shapes. As a consequence, the Projection Principle cannot rule out these deviant scenes as possible interpretations of Picture E. And if such non-cube shapes cannot be ruled out by the Projection Principle, then the principle alone cannot explain why the property of *being a cube* seems to figure in E's content. For the same reasons, the principle alone cannot explain the attribution of high-level properties, like *being a person* or *being a plant*, nor even relatively low-level spatial features like depth, shape, or texture. Yet at least some of these are clearly part of a picture's content.



The fact that there are aspects of pictorial content that go beyond those mandated by the Projection Principle is the reason why it imposes only a *necessary* condition depiction, but does not aim for sufficiency. This point distinguishes the present articulation of projection semantics from that of Greenberg (2013, 249-253), who makes projection both necessary *and* sufficient for accuracy. As a consequence, a picture's content is equated with its skeletal content. But this result is unpalatable, for it rules out obvious features of pictorial content like depth and shape, as well as any kind of high-level attribute (Sober 1976, 113).

In addition to the spatial content specified by a projection semantics, a final theory of depiction must also account for those aspects of pictorial content which have been the traditional focus of "depiction-as" approaches. These include the **singular content** of pictures– what they are *of* or *about*— as well as their high-level **attributive content** — the category-level properties and relations they attribute, beyond the low-level spatial features discussed here (Greenberg 2018, 866-870).

Singular content includes the concrete individuals a picture may depict, but also those non-existent, fictional, and generic objects which may figure in pictorial content. A projection-based semantics is compatible with a variety of approaches to singular content, including those which foreground causal context, artist intentions, or recognitional potential, so long as they integrate pictorial objects into an overall pictorial space.[20] Meanwhile, high-level attributive content seems to stem from a range of general capacities or principles that go beyond systems of depiction, potentially involving both perceptual engagement and pragmatic rationality (Kulvicki 2006, ch. 9).[21] Such mechanisms must, in Kulvicki's words, "flesh-out" the skeletal content supplied by a projection semantics.

Thus, in the kind of final story envisioned here, a picture's content is the nexus of many factors, including the artist's contextually-situated intentions, general perceptual and pragmatic capacities, and systems of depiction grounded in projection. Pictures are mobilized to engage all of these mechanisms and channel them through a coordinated visual frame. What I've aimed to do here is describe the underpinnings of this complex representational feat.

3 From Projection to Pictorial Space

The thesis of this essay is that projection semantics provides the best account of the foundations of pictorial space, for a central class of pictorial representations. Specifically, I'll argue in this section that the Projection Principle explains the mapping from pictures to the *directional organization of basic features* in pictorial space, and that this organization in turn grounds any further elaboration of pictorial space.

At the outset I defined a pictorial space as a three-dimensional arrangement of objects and properties relative to a viewpoint. The location of each entity in a pictorial space is specified in terms of two parameters. The first is the *direction* of that entity relative to the viewpoint, and the second is the *depth* of that entity *from* the viewpoint. As I'll show, the Projection Principle is silent on matters of depth, but it does predict the full suite of directions that make up the substructure of pictorial space.

The Projection Principle fixes not just a set of directions, but also a set of low-level features that may be located in each direction. Here we may draw a distinction between *basic features* and *high-level features* within pictorial content. In a given system of depiction, we may define the **basic marks** as the minimal set of marking types from which all pictures in that system may be generated. In a simple system of black and white drawing, for example, the basic marks are black lines and white areas, or perhaps black points and white points. A **basic feature** is a property that is expressed by a basic mark no matter the visual context. For example, in the simple line drawing system, whether I use lines to draw a tree, or a person, or a cube, black lines will always be used to depict edges, and white patches to depict surfaces. Thus

[20] For example, see Greenberg (2018, 867-70, 893-94) for an account of singular and generic content which is compatible with a scene-based approach to a picture's overall content.

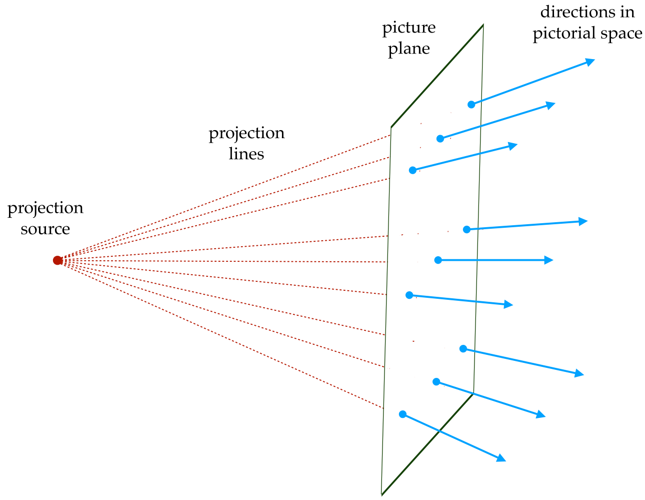
[21] The need for levels of content beyond the skeletal are vividly demonstrated by the case of ambiguous figures (Voltolini 2015, 42-43). Voltolini further argues that the relevant perceptual capacities which help flesh-out ambiguous figures also operate at the level of recognition of the pictorial vehicle itself (pp. 127-31). Though I have provisionally treated pictures as syntactically unstructured here, I welcome the more structured account that these observations suggest.

the basic features of this system are *edge* and *surface*. Other systems will contribute other basic features: in a system of color depiction, for example, the basic-features will be a possibly infinite set of color-like properties.

A complication is introduced by the flexible use of markings in more naturalistic systems, like drawing systems which make use of lines in different ways. For example, lines can be used both to depict edges and to form shading gradients. Indeed, a marking condition like this is at work in Picture A. In a flat-footed description of such systems, the basic features expressed by a line, say, will be highly disjunctive, something like: *edge or shaded region*. They nonetheless express basic features in the sense defined here. A more illuminating approach is to count the different kinds of line as different kinds of basic mark, even if they have the same local physical realization. The result is that the basic marks can now be assigned more natural basic features; one kind of line express *edge* and the other *shaded region*.

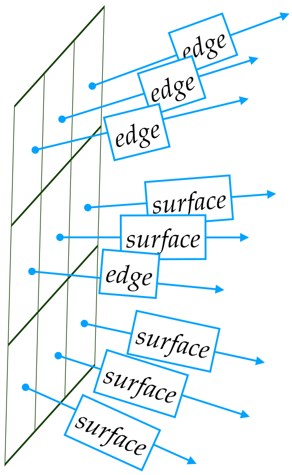
As I'll show, the Projection Principle associates every picture in a given system with the direction of basic features in its content. To see why such a mapping is implied, we must work backwards. The principle states that any picture in a system is a projection of the space it expresses. Thus, given a picture, we determine the kind of space it must express by inferring from the picture back to the location and features of the objects that must have projected it.

For example, if a black and white line drawing contains a black line at point *p*, we can infer that *p* is projected from an unoccluded edge, rather than a surface, because it is black rather than white. And we can further infer that this edge is located along the projection line that intersects the viewpoint and *p*. Thus the picture's content locates an edge in that direction. The same reasoning can be carried out for every point in the picture, assigning edges and surfaces to each direction radiating out from the viewpoint. The figure below illustrates the relationship between projection lines and directions in pictorial space, applied to points across the picture plane.



Thus, by the Projection Principle alone, a picture expresses a content which specifies the directional location of basic features relative to a viewpoint, for every point on the picture plane. This, ultimately, is the spatial import of skeletal content and the contribution of projection semantics. Projection semantics can now be understood not merely as an abstract constraint on content, but as the source of the most basic layer of pictorial space.

The content at work here is usefully conceptualized (and visualized) as a type of feature map where each point in the map is associated with a pair of a direction and a basic feature (c.f. Marr 1982, ch. 4, Treisman 1988, Tye 2000, ch. 5).[22] The result can be visualized by "extruding" the directions, below. This **perspectival feature map** is a kind of semantic convolution of the picture— a metaphysical drawing derived from the original by replacing lines and white space with edges and surfaces, and 2D locations with 3D directions. Such a map is accurate when, for each point in the array, the associated basic feature is located at some point in the associated direction.



The skeletal content of a picture— that part of a picture's content contributed exclusively by the picture under projection semantics— corresponds exactly to the content of a perspectival feature map populated with directions and basic features. As this representation makes clear, skeletal content makes no commitment about depth or high-level properties, but is entirely determinate about directions of basic features. This minimal structure may subsequently be enhanced with objects, properties, and relations by associating them with different regions of the feature map— but each of these will perforce conform with the base-line directional organization.

Whatever further content a picture expresses, it only does so (partly) in virtue of locating basic features in directional space. Picture A may depict a figure in the foreground as exhibiting the property *sitting down*. But it only does so in virtue of

[22]The Projection Principle determines a unique perspectival feature map for every picture, on the assumption that, in the content of each picture, the internal structure of the general viewpoint— the relationship between projection source and plane— is held fixed across scenes. This assumption accords with intuition, because there does not seem to be ambiguity about the directional structure of space in a given picture, even when there is ambiguity about depth and shape. If projection source and picture plane are allowed to vary with respect to one another among scenes within the content of a single picture, then the Projection Principle associates each picture with a set of perspectival feature maps, rather than a unique map.

depicting their legs and torso in a particular position relative to the ground. And this only in virtue of depicting each surface and edge belonging to their body in particular locations in pictorial space, and likewise for the ground. A chain of dependence invariably leads from ascriptions of higher-level content to the core commitments of skeletal content, the directional location of basic features.[23] What's more, skeletal content seems to be the *most* fundamental level of pictorial content, since there is no additional content that must be expressed in order for skeletal content to be expressed. This makes the projective analysis a natural starting point for pictorial semantics.

The forgoing analysis makes explicit one sense in which pictorial representation is a form of **iconic** representation, together with maps, diagrams, and 3D models, among others. Iconic representation is widely thought to trade upon distinctive natural correspondences between the structure of the representation itself and the structure of the content it expresses (Giardino and Greenberg 2015, 3-4). In Venn and Euler diagrams, for example, these correspondences take the form of isomorphisms between space on the page and logical space (Shin 1994, 170-73; Hammer and Shin 1998, 15-18); in iconic representations of magnitude they may take the form of multiplicative or logarithmic transformations between representation and quantity (Beck 2015, 8-10). In the present analysis, projection itself is the natural relation which binds together a pictorial representation and its content. While pictures express content well beyond skeletal content, their iconicity is most pronounced in the relation of the layout of the pictorial surface to the directional layout of the pictorial space it expresses.

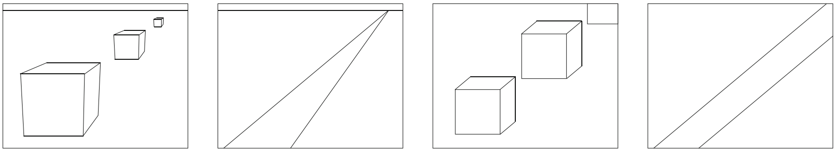
4 Systems of Depiction

Any answer to the problem of pictorial space must accommodate the central fact that the expression of pictorial content is always relativized to a **system of depiction**, the pictorial analogue of a language (Gombrich 1960; Hagen 1986; Willats 1997). Systems of depiction vary along a host of dimensions, including their treatment of geometry, line, light, and color. These variations give rise to pictorial spaces with different directional structures, made up of different basic features. In this section I'll show how the Projection Principle can be extended to explain the foundations of pictorial space for a broad class of systems of depiction.

Differences in perspective geometry are vividly illustrated by the contrast between **linear perspective** (F and G) and **parallel** (H and I) systems of depiction. Pictures in linear perspective marked by the fact that objects at greater distances from the viewpoint are depicted by smaller regions on the picture plane (F), and parallel lines receding from the viewpoint are depicted by converging lines on the picture plane (G). By contrast, in parallel systems, even as objects move farther from the viewpoint,

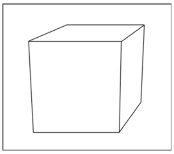
[23]The dependence of high-level content on skeletal content is asymmetrical. In general, for any given bit of high-level content *C* (which may be only a part of a picture's total content), it appears to be impossible for a picture to express *C* without also expressing some skeletal content. The reverse is not true, because there are genuine skeletal contents– those expressed, for example, by fields of static or indecipherable marks– that lack any substantive high-level content.

they are depicted by regions of the same size on the picture plane (H); and parallel lines receding from the viewpoint are depicted by parallel lines on the picture plane (I). Parallel systems give rise to a kind of unsituated "god's eye view," though such pictures are still perspectival, since they clearly depict their subjects from a particular direction.



F G H I

While linear perspective has ancient roots, it was first codified during the Italian Renaissance, and now dominates contemporary mass media.[24] Meanwhile, parallel systems are the norm in architecture, engineering, and other technical fields; they are used for maps of limited geographic extent; and they are commonplace in the art of classical East Asia and ancient Egypt. There are many types of parallel system, including isometric, axonometric, and orthogonal projection systems (Dubery and Willats 1972).



J

Pictorial systems like these are not merely *styles*— guidelines for construction with no semantic consequences. Rather, they embody genuine rules of interpretation which associate different kinds of pictorial space with the pictures in their domains. On their face, pictures in perspective seem to express spaces that are organized radially around a point of view, whereas pictures in parallel systems seem to describe a comparatively flat and uniform layout. To make the difference concrete, suppose we isolate Picture J without specifying the intended system of depiction, and ask after its content. Relative to the linear perspective system, it may express a space containing a cube, but relative to a parallel system, it can only express a space containing an irregular solid. This is because, in linear perspective, the converging lines on the picture plane may be interpreted as representing the parallel edges of a cube; but in parallel systems, converging lines can only be interpreted as representing converging edges. Thus different systems of depiction determine different ascriptions of content, in part by associating pictures with spaces of different structure.[25]

[24] Linear perspective is often simply called "perspective." But it should be distinguished from curvilinear perspective, the sort of image produced by a fisheye lens.

[25] The same point can be made with respect to the difference between systems of black and white line drawing and systems of color line drawing. Relative to a system of black and white line drawing, Picture

Normally, however, the operative system of depiction is relatively well fixed, even if its identity is not fully known to all viewers. The selection of operative system of depiction results from an interplay between artist's intentions and local convention. In cases where local convention is especially strong— as in thousands-years-old artistic traditions, or in specific types of media, like map books or newspapers— convention may dominate. In situations where local convention is more open-ended— as when sketching a landmark as part of giving detailed directions— intention may select among, or even extend, the available systems. In that case, interpreting the picture is at least partly a process of recognizing the system-directed intentions of the artist.

According to the Projection Principle, systems of depiction affect content by fixing the underlying method of projection. To interpret a picture as part of a perspective *system* is to treat it as if it were produced from the corresponding *method of projection*. I provisionally assume that every system of depiction determines a unique method of projection, what I will call its **characteristic** method of projection.[26] The spatial variation between parallel and linear perspective systems can be attributed to alternation between underlying methods of projection.

The method of projection that was featured in previous sections is known as **linear perspective projection** (for short, **perspective projection**); it is characterized by the fact that the projection source is a *point*, to which all projection lines converge. An alternative, known as **parallel projection** replaces the point-sized projection source with an oriented *plane*.[27] Then the projection lines, rather than converging on a single point, all run perpendicular to the projection source, hence parallel to one another. An example is illustrated below. In this case, the projection source and picture plane are themselves parallel, but this is not required. Even though the projection source of parallel projection is a plane, I still refer to the combination of projection source and picture plane as a *viewpoint* since it provides a type of spatial perspective on its subject matter.[28]

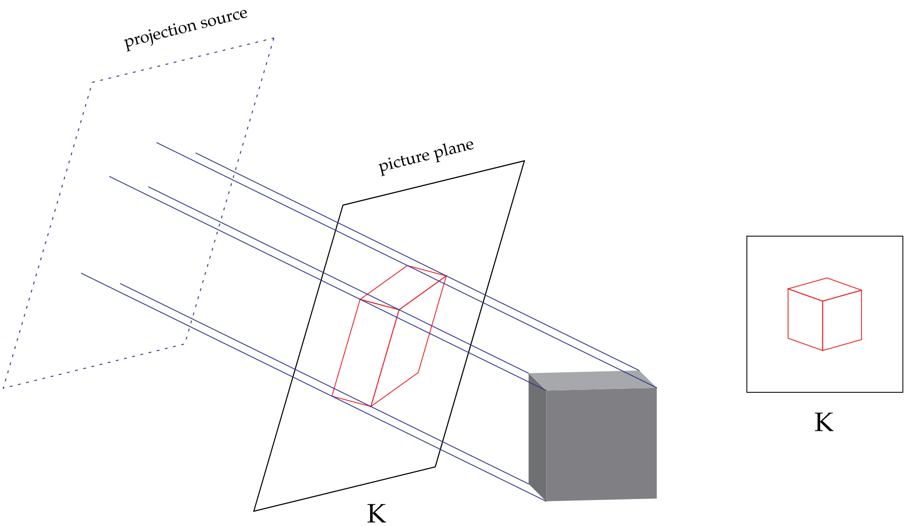
J might depict a solid of any color or shade— its content is indeterminate in this respect. Relative to a system of color line drawing, it must depict a white solid on a white background.

[26]This isn't always the case. There are systems of depiction which involve more than one method of projection. Generally such systems fall under the rubric of "impure projection," discussed in Section 7.

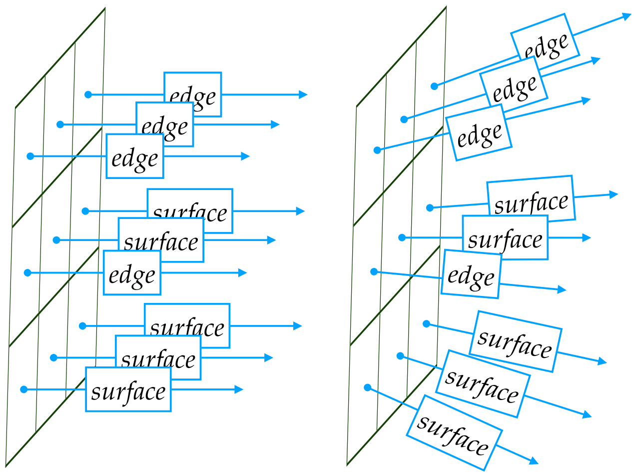
[27]Hagen (1986, 142-43) offers an alternative analysis, according to which the projection source of parallel projection is actually a *point*, but it is located at infinite distance from the scene. But this analysis is unwieldy to implement in geometrical terms; and it undermines any finite attributions of depth that might figure in a picture's fleshed-out content.

[28]The familiar examples of axonometric, isometric, and elevation projections are all species of parallel projection, derived in part by varying the relationship between projection source and picture plane. Still other methods of projection can be defined by varying the structures of the projection source, picture plane, projection lines, and their relative relationships.

In the text I discuss methods of projection where the projection lines are themselves straight lines. But in a prominent alternative class of projections, lines of projection instead closely recapitulate the behavior of light rays, reflecting and refracting in the environment before reaching the picture plane. Such optical methods of projection are the norm in realist painting, computer animation, and, of necessity, photography. The construction of images in modern 3D animation uses a method of rendering known as *ray tracing*— essentially tracing the complex path that light would realistically take in the scene.

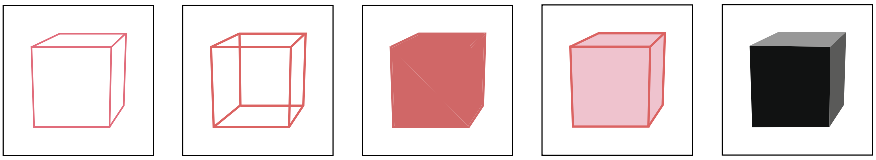


The differences between perspective and parallel methods of projection result in the interpretive variations noted earlier, for example, the significance of converging as opposed to parallel lines on the picture plane. The underlying explanation of these facts is that pictures in the two systems express pictorial spaces which themselves have different structures, in particular different *directional* structures. By the same reasoning introduced in the last section, we may derive a directional feature map from the structure of projection lines in a parallel projection system. In this case, the specified directions are all parallel to another, rather than converging backwards to a viewpoint. The resulting structure is illustrated at left below, set beside the feature map for linear perspective at right.



Besides perspective and parallel projections, other forms of projection entirely include curvilinear perspective, reverse perspective, and globe projections. The latter, employed for maps, variously place the projection source above, below, or even *within* a sphere, in order to project it onto a conic, cylindrical, or flat picture plane (Robinson and et al 1995). Thus the projective analysis confirms the proposal, set out in Section 1, that most maps form a natural class with what are traditionally counted as pictorial representations.[29]

As we saw in Section 2, methods of projection are defined by a combination of a projection condition and marking condition (Willats 1997, 4-20). Where the projection condition determines the directional structure of pictorial space, it is the marking condition which fixes the basic features which tile this space. Variation in the marking condition gives rise to a diversity of treatments of line, color, and shading. Abstractly, the marking condition has the function of defining mappings from types of features in the scene to types of marks on the picture plane, which are then applied to the spatial distribution specified by the projection condition. Such an analysis can be applied to color paintings, photographs, X-rays, drawing with shading, and many different kinds of line drawing.[30] The figure below illustrates variation in markingconditions.



edges → lines surfaces → white | *all edges → lines* | *surfaces → red* | *edges → lines surfaces → pink* | *colors → colors*

Through variations in the projection and marking conditions, the concept of projection, and thus the Projection Principle itself, can be extended to a surprisingly wide range of systems of depiction. Bringing together research programs in art history, computer vision and graphics, and technical drawing, scholars like (Hagen 1986; Willats 1997; Durand 2002), and others, have shown that an impressive range of art historical "styles" can be analyzed in terms of the basic ingredients of projection, well beyond those canvased here.

Among the methods of projection covered by the Projection Principle, those which diverge from the behavior of light projected to the eye are of special significance. Following Willats (1997, ch. 6), we may roughly distinguish optical and non-optical

[29]As an anonymous reviewer points out, topological (or quasi-topological) maps, such as subway maps, as well as circuit diagrams, do not fit this mold. Still, they may be derived from projections by abstracting away from metric properties; I'm inclined to think of these as a blend of pictorial and diagrammatic representations, exploiting interpretive competencies from both visual and geographic cognitive modalities.

[30]There has been extensive research on the marking conditions at work in line drawing: see e.g. Kennedy (1974, chs. 7-8), Willats (1997, ch. 5), Palmer (1999, §5.5.7), DeCarlo and et al (2003). For discussion of color and other "optical" methods see Willats (1997, ch. 6), who shows that even photographs can be conceived as the products of projection. Most photographs exploit a projective geometry approximately like linear perspective with ray-tracing, and an optical system of marking loosely analogous to the "color-to-color" method illustrated in the text.

methods of projection. **Optical** methods of projection cleave as much as possible to the physical behavior of light: in terms of the geometry of the projection lines and in the encoding of color and shading.[31] Such projections are characteristic of realist, linear perspective painting styles, and many forms of color photography; they produce the kinds of images which, under suitable circumstances, can become *trompe l'oeil*. But there is a proliferation of **non-optical** methods of projection as well. Many have argued that linear perspective itself, though similar to optical projection, diverges from its fine geometry (Pirenne 1970; Hansen 1973; Rogers and Rogers 2009). As for marking, even black and white photography is a departure from optical projection. More dramatically, systems of line drawing involve the registration of edges as lines, quite unlike the information registered by the retina. Correspondingly, no picture resulting from a non-optical projection could serve as a *trompe l'oeil* depiction of its content.

By allowing in the full range of systems of depiction, optical or otherwise, the Projection Principle casts a wider net than many of its its progenitors. Hyman's (2006, ch. 5, 2012) principle, for example, is keyed to only a narrow range of nearly optical systems, while authors in the perceptual resemblance tradition, including Peacocke (1987), Budd (1996), and Hopkins (1998), discussed in Section 6, tether their theories directly to the geometry of optical projection in perception.[32]

Kulvicki's (2006) account covers a more varied range, but still confines its scope to those systems which satisfy certain formal properties, constraints which do not apply to the Projection Principle.[33]

Yet even the comparatively flexible projective account offered here has its limits. In Section 7, I'll turn to the case of stylized depiction, which resists a strict projective analysis, but, I will argue, is nevertheless constrained by projection in a looser

[31] In terms of the geometry of projection lines, optical methods are characterized both by a perspectival organization of projection lines, and by a ray-tracing approach to their propagation through space. Systems can be optical with respect to geometry, but not markings, and vice versa.

[32] Hyman (2006) writes as if projected shape (for Hyman, "occlusion shape") can always be defined relative to a point-sized projection source (p. 76), with projection lines that follow the behavior of light (p. 77). In addition, since Hyman's Occlusion Shape Principle is modeled after optical projection to the eye, it doesn't distinguish the elements of picture plane and projection source, which in the present framework make up a viewpoint. As a consequence, the principle fails for systems of *anamorphosis*, which involve projections where the projection source and picture plane are positioned at an acute angle with respect to one another. Since the Occlusion Shape Principle is invalid here, Hyman (2006, 93-98) is compelled to introduce an additional principle for the special case of anamorphosis. By contrast, the Projection Principle handles anamorphosis as a yet another system of depiction, albeit one which imposes distinctive constraints on viewpoint.

[33] To be precise, Kulvicki holds that for a system to be pictorial, it must exhibit "transparency" *or near enough*; see ch. 3, especially p. 76. Intuitively, transparency requires that a picture *of* a picture have the same syntactic structure as the original. In Kulvicki's system, this amounts to the claim that pictures in a given system satisfy their own skeletal content. (See the Appendix for a more detailed discussion.) Suffice to note for now that perspective and parallel projections are transparent, while curvilinear systems are not (Greenberg 2013). Whether this marks a contrast with the Projection Principle depends on the scope of Kulvicki's "near enough" clause. More vivid challenges are marking conditions used in edge- and topography-based line drawing. Such systems are stark violations of transparency, but standard cases for the Projection Principle.

sense. For now, my central claims are confined to the core class of "realistic" and mechanical pictorial systems.

5 Projection and Vision

I have argued that the expression of pictorial space can only be explained with the help of a projection semantics. But pictorial representation is, at the end of the day, a *human* activity, and systems of depiction are types of signaling system carried out with human artifacts. So why would humans employ a signaling system that is based on projection in this way?

It is clear that projection is, at some level, an abstraction and idealization of the optical projection of light reflected to the eye. Just as light carries information about surfaces that it has reflected off of to receptor cells in the retina, lines of projection transfer features of the scene to features of the picture plane.[34] But this analogy, while apt, only illustrates the commonality between projection, construed as a method of picture production, and optical projection, a process for producing a retinal image. Yet, to reiterate a theme, the Projection Principle is a principle of *interpretation*, not of production. The correspondence between geometrical and optical projection does not yet explain why geometrical projection is also recruited as a semantic constraint on pictorial content.

Instead, I propose that the Projection Principle is not directly grounded in optical laws of projected light, but rather in the computational commitments *about* those optical laws encoded in the visual system. According to the computationalist understanding of vision, a central function of the visual system is to generate an estimate of the kind of scene that must have produced the retinal image (Marr 1982, 32-38). But here vision faces the "inverse problem:" for any given image, infinitely many scenes could have caused it, while only one scene is the true cause. The visual system overcomes this problem by making a best guess, informed by implicit heuristic assumptions about the environment, and about how the environment influences the retinal image (Palmer 1999, 80-85). The most basic such assumption is that the retinal image itself is the product of the optical projection of light through the lens of the eye (Rock 1983, 324).[35]

Without this background constraint, further assumptions about regularities in the environment cannot be reliably related to the retinal image. Thus encoding the laws of optical projection is the ground upon which more refined visual computations must be built.

[34] To spell out the analogy: the retinal surface is analogous to the picture plane; the focal point defined by the lens of the eye is analogous to the projection source; and the ray-like behavior of light which links the scene to patterns of activity in the retina resembles the straight lines of projection.

[35] The sense in which such laws are "assumed" by the visual system is an open question; it is unlikely that they are actually *represented*, but they may well be encoded in the normal operation of the system itself (Johnson 2020, 9-13). Either way, the inverse problem is intractable unless assumptions about the projective behavior of light are in some way respected.

My conjecture is that the Projection Principle is recruited as a principle of interpretation because the visual system already uses an analogous principle in its normal operation. On one hand, the visual system hypothesizes an environmental space on the basis of a retinal image. On the other, a viewer hypothesizes a pictorial space on the basis of a pictorial image. In the first instance, the hypothesis is an (unconscious) inference to the best explanation about the actual environment. In the second, the hypothesis is an interpretation, a guess about the type of scene the image purports to be projected from. The Projection Principle thrives in human transaction because the computations it requires can be carried out on much of the same computational machinery already supplied by the visual system. This does not mean that pictorial interpretation and perception are the same, only that they rely on analogous rules, and recruit overlapping cognitive capacities.

Meanwhile, from the perspective of human communication, the enlistment of vision as a mechanism for interpreting signals seems nearly inevitable. The human visual systems is computationally powerful; all normal humans have one; and all normal human visual systems are very similar to one another. A signaling system which exploited these facts would have at its disposal a fast and reliable interpretive engine shared by most interlocutors. This alone would obviate much of the individual work typically involved in learning a communicative code, and much the collective work of coordinating on one. As a consequence, in so far as methods of projection can be directly computed by the visual system itself, they will naturally be favored for use in communication. But in so far as pictorial representation is subject to external constraints, it may stray from the strictures of vision. These pressures may include limitations of artistic skill, ease of construction and ease of use, communicative or perceptual effect, and the need for standardization (Giardino and Greenberg 2015, §1.2). Systems of depiction are in part the product of this conflux of innate visual ability, practical demands, and the interpersonal drive to communicate.[36]

Still, there are limits on the kinds of possible signaling system which my reasonably count as pictorial. The argument of this essay suggests that projection in some form grounds a broad class of central cases of pictorial expression. But if projection need not be optical, how far can it stray from its optical origins? The problem is that there are “methods of projection,” broadly speaking, which are clearly non-pictorial. I have characterized projection as a mapping from a three-dimensional scene to a two-dimensional picture plane via a systematic set of spatial relations defined by a viewpoint. And I have given concrete examples of core cases. But one can imagine a spectrum of cases, deviating from this core but adhering to the general description, which seem less and less pictorial at each step. Consider projections which bend

[36]Communication, in the relevant sense, involves inter-personal coordination of information, and isn’t limited to the transfer of information. Thus an architectural plan might facilitate coordination among a group of builders discussing the best course of action, even when all parties were antecedently familiar with the plan. As an anonymous reviewer suggests, points parallel to those made in the text could also be made for *intra*-personal coordination. The stability and power of the visual system makes it possible for a single individual to exploit pictorial signs to extend their access to visual information over time, and in doing so, their capacities for visual reasoning.

and distort their subject matter in unusual ways, in the manner of a fun-house mirror. Such images seem essentially pictorial, albeit in a somewhat degraded way. But more extreme cases resist inclusion: consider a mapping from scenes to pictures that exploits a specific configuration of seemingly scrambled projection lines. The results of such a projection would look like static or confusing swirls of color. Still, it could be a "projection": the viewpoint for such a method might take the form of a highly fractured surface; and it might use the same scrambled configuration for every picture, hence qualifying as a "systematic" spatial relation. Nonetheless, the products of such a method would not qualify as "depictions."

I propose that the category of systems of depiction may be defined in part through functional associations with vision, even though systems of depiction cannot, in general, be understood directly in terms of optical projection or the normal operation of the visual system.[37]

Within cognitive science, the visual system is understood to be a highly modular information processor. Different subsystems handle a wide range of distinct visual computations, including the perception of shape, motion, depth, color and so on. My hypothesis is that, to be pictorial, a system of depiction must be **modularly derived** from assumptions already encoded in the visual system. Such derivation may involve selectively modifying the parameters that control certain visual modules, and it may involve suppressing other modules altogether.[38]

For example, black and white line drawing seems to make use of the visual system's tendency to interpret abrupt color boundaries (lines) in the retinal image as *edges*— but it selectively suppresses the visual system's disposition to interpret regions of color in the retinal image as colored *surfaces* in the environment. This makes sense in the framework of modular derivation, because shape boundaries and color are thought to be computed by distinct modular subsystems (Frisby and Stone 2010, ch. 10).

By comparison, different projective geometries may be derived, not by ignoring the visual system's native projection condition, but by tuning the parameters it depends on. Parallel projection and perspective projection lie on a spectrum, derived by systematically adjusting the angles at which projection lines intersect the projection plane. Indeed, modulation along this dimension may occur naturally in vision, when visual focus is shifted from objects in the foreground (which appear approximately in perspective projection) to objects in the deep distance, as when looking

[37] The theoretical situation is akin to one familiar within the philosophy of language. Various formal-logical definitions of the concept of *language* are available, but many of these formal systems are so abstract, complex, or bizarre that they resist the moniker "language" in any straightforward sense. Theorists have tended to respond (e.g. Chomsky 1965, Lewis 1975) by identifying some subset of these formal systems as the humanly usable languages.

[38] My proposal here bears some affinities to that of Kulvicki (2006, 76), since we both identify systems of depiction in terms of their relative proximity to a core class. But Kulvicki defines that class in terms of the formal property of transparency, while I define mine in terms of visual cognition. (See footnote 33 and the Appendix for more on transparency.) The two ideas come apart because systems can vary their proximity to transparency, without varying in their proximity to visual cognition, and vice versa. Systems like edge- and topography-based line drawing, for example, strongly violate transparency, but are closely related to visual cognition.

out a plane window (which appear more nearly in parallel projection). Systems of depiction based on perspective and parallel projection are, according the present hypothesis, the result of fixing the values of this adjustable parameter.

The common theme among such modular derivations from vision is the location of basic features in directional space, hence the implicit assumption of a method of projection. In essence, the range of possible systems of depiction corresponds to those projection-based systems that harness aspects of visual cognition that can be isolated, modulated, and abstracted. The result is a diversity of possible pictorial spaces, all of which share some common thread with the familiar perceptual space of human vision.

6 Alternative Accounts of Pictorial Space

Aside from projection-based approaches, contemporary theories of depiction tend to cluster around two main ideas. One, the impetus for *resemblance theories*, proposes that depiction is grounded in some form of similarity or resemblance between picture and content. The other, the starting point of *perceptual theories*, holds that depiction is grounded in the contentful perceptual responses that pictures elicit. *Perceptual resemblance theories* are a hybrid of the two.[39]

In this section I'll consider how these theories fare with respect to the expression of pictorial space, in particular its directional structure. While proposals in these traditions tend to focus their attention on high-level "depiction-as" ascriptions of content, these same theories often do end up making predictions for low-level organization of pictorial space. In what follows I'll argue that projection semantics provides a better account of directional structure than either resemblance or perceptual theories. The common theme is that these alternative theories either make false predictions about directional structure or are silent on the matter of directional structure. In the first case, I propose to replace the theories in question with a projection semantics; in the second, the accounts may be partially correct, but only insofar as they are confined to high-level pictorial content, and supplemented by a projection semantics below. Either way, projection semantics seems to provide the best explanation for the foundations of pictorial space, and should be counted as a necessary component of a complete pictorial semantics.

6.1 Resemblance Theories

The class of "resemblance theories" of depiction has assumed many forms over the years. At its broadest, it has included any theory that posits a systematic correspondence between the structure of a picture and that of its content— and contrasts with

[39]This taxonomy is not meant to be exhaustive or exclusive. For example, Voltolini's (2015, 16-22, 164-67) recent proposal weaves together elements from all three traditions.

theories which ground depiction primarily in arbitrary conventions or visual perception. Under this conception, projection-based accounts are a species of "resemblance theory" as well.[40]

Still, there are important distinctions to be drawn among different ways of characterizing structural correspondence. The projection-based approach should be contrasted with the enduring idea that depiction is grounded in relations of restricted *similarity* or *isomoprhism* (French 2003; Greenberg 2013).[41] In what follows, I'll reserve the term **resemblance theory** for accounts of this kind.[42] Resemblance theories hold that the content expressed by a picture is both limited by, and in large part explained by, similarities between the picture and elements of its content.

Some, like Peacocke (1987), Budd (1996), and Hopkins (1998), have focussed specifically on experienced similarities between states of *looking* at a picture, on one hand, and those of *looking* at the scene it represents, on the other. I discuss issues specific to this implementation of resemblance theory in the next subsection.

Others, like Neander (1987), Abell 2009, and Blumson (2014), postulate more direct similarities between pictures and their contents, while allowing that the relevant dimensions of similarity may vary with contextual elements such as salience or intention. Even more than perceptual resemblance accounts, these recent contextualist views are especially geared to ascriptions of high-level content. So construed, this approach constitutes an incomplete account of pictorial content. It is notable that, to my knowledge, none of the articulations of contextual resemblance have offered specific proposals about the types of resemblance that might explain low-level facts like directional structure, comparable to a projection semantics. But no conflict need arise: it may be that projection semantics explains certain low-level aspects of pictorial content, and resemblance, other high-level aspects.

Still, it might be wondered whether contextually restricted resemblance could, at least in principle, explain the full fundamentals of pictorial space. Addressing this question, Greenberg (2013, 271-84) has argued that resemblance theories fail to give the correct adjudication of accuracy for images in a range of **transformative** systems of depiction that include curvilinear perspective, globe projections, systems of shifted color (like technicolor or brown-tinting), and contour-drawing. On Greenberg's construal, resemblance theories make similarities between a picture and a scene a necessary condition on the picture being accurate at that scene. But, as he shows,

[40]Hyman (2006, 111) describes his projection theory as "the defensible residue of the resemblance theory of depiction."

[41]In the latter case, isomorphism is understood as a kind of abstract structural "match" between two spatially defined systems. Despite its apparent abstraction, isomorphism of this kind is just another form of similarity— similarity with respect to abstract, relational features.

[42]There are other kinds of account which might also be called "resemblance theories," but which which cast the concept of similarity in quite different roles. One "meta-semantic" theory holds that for a system to be pictorial (or perhaps, iconic), similar pictures must express similar contents— hence the space of possible pictures is similar or isomorphic to the space of possible contents. (Thanks to John Carriero and Matthew Fulkerson for suggesting this idea.) The Projection Principle is, as far as I can tell, fully compatible with this proposal.

the systems in question require dimensions of *difference* between pictures and the scenes they are accurate at. Curvilinear perspective imposes differences in geometry between picture and scene, while color-shifted systems require differences of color.

These conclusions shed light on the present discussion because the errors which Greenberg ascribes to resemblance theories— diagnosed as errors in determinations of accuracy— derive from mistakes in the attribution of directional structure of pictorial space. Greenberg's argument implies that resemblance accounts deliver the wrong attributions of directional structure to pictorial space for curvilinear systems and the wrong attributions of basic features for color-shifting or contour drawing systems.

From this we may conclude a disjunction. Either resemblance is defined strictly (as Greenberg imagines) and delivers the wrong attributions of pictorial space to curvilinear, shifted color, and other transformative systems; or resemblance is defined loosely, and is, at best, silent on fine-grained attributions of pictorial space in these cases. (The silence may be committal or not, depending whether resemblance is thought of as a sufficient as well as necessary condition on depiction.)

Meanwhile, projection semantics specifies attributions of low-level spatial content which are both correct and fine-grained for the full range of pictorial systems under discussion, including those of curvilinear perspective, systems of shifted color, and beyond (Greenberg 2013, 264-267). Unlike resemblance, projection is a *transformation*, one that must be defined in terms of both the similarities it preserves and the deformations it introduces. This makes projection semantics uniquely well-suited to capture the expression of skeletal content in systems that require dimensions of difference between picture and content.[43]

In sum, whatever their merits for the analysis of high-level depiction, the coverage of resemblance theories is perforce limited in scope, and constrained in spatial detail. At least with respect to the directional structure and basic features of pictorial space, projection seems to be the more flexible and general notion, and the one that most readily yields precise analyses.

6.2 Perceptual Resemblance Theories

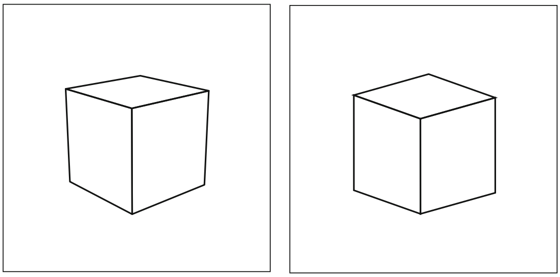
The next class of resemblance theories warrants special attention here, because of the detailed predictions it makes about pictorial space, and the central role it assigns to projection. Developed by Peacocke (1987) and Budd (1996) and Hopkins

[43]The distinction between resemblance- and projection-based theories is sometimes clouded because projective invariants are superficially analogous to the properties held fixed in a resemblance theory. To be sure, projection does entail certain projective invariants— properties of a scene that always survive projection to the picture plane. But, in general, the relation of projection cannot be equated with that of sharing projective invariants, because nearly every form of projection imposes systematic differences between its relata that go beyond invariants. In addition, projective invariants do not even imply shared properties between pictures and scenes. For example, the property of line straightness is invariant under linear perspective projection: if an edge is straight in the scene, its projection will also be straight. Yet the reverse is not true. For there can be straight lines in the picture which are *not* projected from straight edges in the scene, but instead from curved edges rotated perpendicularly to the picture plane. So while line straightness is an invariant under perspective projection, it is not in general a property shared by a picture and its content. Thus projective invariants provide little help for resemblance theories.

(1998, 2003, 2006), I'll collectively call these **perceptual resemblance theories**. For concreteness, I'll generally follow Hopkins' articulation of the underlying idea. Perceptual resemblance theories pursue the common thought that, for a picture to depict an object as having a given property, the picture and an object with that property must be taken to *look alike* in relevant respects. Further, *how something looks*, at least with respect to shape, is defined in terms of the shape it takes up in the phenomenal visual field; it is further assumed that an object's visual field shape standardly corresponds to an optical projection of the scene before the viewer. Hopkins (1998) calls this the **outline shape** of an object.[44]

According to perceptual resemblance theory, the content of a picture is the type of scenario that the picture is experienced as resembling in outline shape. (To fix content, the experience must also satisfy certain correctness conditions, but I set these aside for now.) A suitable viewer, when looking at a picture P, will have an experience with content of the form: *P is similar in outline shape to an F*.[45] The content of the picture is thereby a scenario that includes the property F. So if a suitable viewer would experience a picture as resembling a *horse* in outline shape, then the content of the picture will include the property of being a horse. Note that, for perceptual resemblance theories, it is only *optical* projection which defines outline shape; there is no in-built relativity to other systems of depiction, as there is in projection semantics. It is this feature which is the focus of my criticism here.

In this section, I'll argue that systems of depiction with non-optical geometry pose a fundamental challenge to perceptual resemblance theories, even while projection semantics accommodates such cases smoothly. Hopkins himself has outlined various ways of addressing this challenge, but I will argue that these replies are inadequate to the problem at hand (Hopkins 1998, ch. 5, 2003, §4-6, 2006; Lopes 2006). To see the basic issue, consider L, an accurate picture of a cube in a system of linear perspective, and M, an accurate picture of a cube in a system of parallel projection.



L M

[44]This idea is expressed differently by the different authors. Peacocke (1987) writes of *shapes of regions in the visual field* (p. 385). Budd (1996) uses the idea of *two-dimensional aspects of the visual field* (pp.158-159), understood as a two-dimensional abstraction of three-dimensional perceptual space. And Hopkins (1998, ch. 3) defines *outline shape* as the solid angle subtended by the object relative to a viewpoint. This is not quite the same as a shape on the visual field, but the difference is inconsequential, since every visual field shape determines a unique outline shape, and vice versa.

[45]Note that the relevant experience of similarity is between the outline shape of P and the outline shape of an F— not between P (simpliciter) and the outline shape of an F.

Under the right viewing conditions, the outline shape of the picture L and the outline shape of a cube would be the same. Thus it would be normal for a viewer to experience L as resembling a cube in outline shape. Since this is what the picture accurately depicts, the perceptual resemblance theory goes through without a hitch. But now consider M. The outline shape of M is noticeably *not* the outline shape produced by looking at a cube. A standard viewer would not normally experience M as resembling a cube in outline shape. How can the perceptual resemblance theorist reconcile this fact with the starting assumption that M is an accurate depiction of a cube, in the system of parallel projection?

A possible reply is to deny the starting assumption that the a picture like M could ever be an accurate depiction of a cube. But this would imply that nearly all pictures in parallel systems misrepresent their subjects. One problem is that parallel projection in particular is the norm in technical and engineering drawing, where the accuracy of such pictures is assumed and relied upon. The claim of nearly universal misrepresentation seems to contradict the facts on the ground. A second issue is that, counter to intuition, we would no longer be able to distinguish between accurate and inaccurate depictions for pictures in non-perspectival systems (Hopkins 2003, 152).

A different response defended by Hopkins (1998, ch. 6; 2003, §4), is to loosen the demands on content fixation so that the content of the picture may be less determinate than the content delivered by a fine-grained experience of resemblance. In effect, the noticeable differences between the shape experiences of M and of a cube are treated as irrelevant to the determination of content, while the noticeable similarities are foregrounded.[46] The result is that the content of the picture is strictly less determinate than the content that would normally be ascribed by perceptual resemblance. Specifically, it must be significantly imprecise about the directional location of basic features, the orientation and length of edges, and the shapes of edge-bound surfaces. But as a result parallel pictures like M can be counted as accurate. Hopkins himself embraces this conclusion, noting that it generalizes to pictures in all systems which deviate from the optical geometry of perspective: “only picture in perspective depict detailed spatial content” (Hopkins 2003, 162).[47]

The idea that some kinds of pictures necessarily express indeterminate spatial content is plausible for a rarely used system like inverse perspective (Hopkins 1998,

[46] In fact, Hopkins has various ways of reaching this conclusion. One is to allow that the features of the picture plane which are the subject of experienced resemblance are only those which are especially salient; and for some systems, only rough shapes, not fine-grained spatial detail, is salient. The result is that experienced resemblance may be to *a table* or even *a cube*, but not one with narrowly defined spatial dimensions (Hopkins 2003, 153). A second tack is to allow that, under certain circumstances, the actual content of a picture can be less determinate than experienced resemblance (Hopkins (1998, ch. 6). Hopkins (2003, 155) indicates that this is his preferred approach to the problem raised by parallel projection.

[47] Hopkins’ conclusion here is not that pictures in the parallel system are entirely indeterminate or devoid of substantive content. A parallel picture like M may be experienced as resembling a *cube* in outline shape; but it cannot be experienced as resembling anything more specific in outline shape— a cube of certain dimensions, or a cube in a certain orientation, for example. An experience of resemblance to one of these more determinate situations would require a more precise match in outline shape. So the perceptual resemblance theorist can allow that, for pictures in parallel projection, their high-level content may be relatively fleshed-out, even though there would still be considerable indeterminacy about low-level spatial content.

157-8), which is difficult to work with and difficult to interpret. But for systems of parallel projection which are the focus here, the claim of necessary spatial indeterminacy is surprising on its face: the parallel picture M does not seem to convey less exact spatial information than the perspective picture L. Given the instruction to construct a model of the object depicted in each, I would not feel more latitude in the case of picture M than picture L (just the opposite, perhaps). This reaction is reinforced again by the pervasive use of parallel projection in technical and practical contexts. Parallel projection is the norm in architecture, for example, not only because of its potential for accuracy, but also because of its capacity for precision. The relative absence of indeterminacy or ambiguity in parallel projection drawings is what has made them so suitable for high stakes pictorial communication, where precision is prized and miscommunication is costly. The idea that there is widespread spatial imprecision in parallel projection does not seem to fit with core uses of pictorial representation in daily life.

The lesson here is that, by fixing optical projection as the basis for all forms of pictorial representation, perceptual resemblance theories are inevitably led into unpalatable conclusions about the structure of pictorial space for non-optical systems. As far as the representation of spatial detail goes in these systems, projection semantics seems to give more satisfactory account, and the more general account overall.

6.3 Perceptual Theories

I turn finally to compare projection semantics with the family of **perceptual theories**, including the accounts of Schier (1986), Wollheim (1987), and Lopes (1996) and Newall (2011), which ground depiction directly in various aspects of visual perception. Schematically, views of this kind seek to derive the content associated with a picture from the content that the visual system would generate when exposed to the picture under suitable circumstances (e.g. when looking at the picture in normal lighting). For example, Wollheim (1987, 46-51) can be read as holding that for a picture to express a given space as its content, it must be possible for a suitable viewer to *see* that space *in* the picture— where the critical notion of *seeing-in* is defined at least in part by the normal operation of visual perception. Other authors, like Schier (1986, ch. 6) and Lopes (1996, ch. 7), claim that for a picture to attribute a property as part of its content, a suitable viewer must *visually recognize* that property in the picture.

In Section 5, I characterized systems of depiction as those projection-based sign systems which bear certain affinities to human vision. In this respect, the projection semantics defended here resembles perceptual theories of depiction. But the two approaches cast visual perception in different theoretical roles. For traditional perceptual accounts, the role of perception is *semantic*— it helps to explain why individual pictures have they content they do, in virtue of the perceptual contents possessed by an idealized viewer. By contrast, I hold that the role of perception is *meta-semantic*— it helps explain why some rules and not others are selected as the interpretive principles of depiction. At least for the low-level features of pictorial space considered here, I do not draw any direct connection between the contents of pictures and the contents of perceptual states.

The degree of conflict between projection semantics and perceptual theories of depiction depends on which aspects of pictorial content resemblance accounts aspire to explain. Insofar as their focus is the depiction of individuals and high-level properties, as appears to be the case of Schier (1986, ch. 6) and Lopes (1996, ch. 7), then there may be no conflict with a projection semantics, and indeed potential synergies. Then the two are simply aimed at different domains, a view I have sympathy for. But insofar as perceptual theories aim to explain the spatial geometry of pictorial space, as Wollheim's (1987, 46-51) remarks suggest, then they are in tension with projection semantics, and I believe, cannot be right.

The problem is the tight connection between perceptual content and pictorial content that perceptual theories impose. Especially in its early stages, human visual perception is thought to work in a manner that is relatively fixed, both among human perceivers, and within perceivers across contexts. The influence of world knowledge and reasoning on low-level spatial vision is negligible (Pylyshyn 1999, ch. 2; Firestone and Scholl 2016). So it is not clear how a set of fixed visual process could yield the interpretive variability introduced by the many systems of depiction. The challenges are now familiar. For example, early vision normally treats certain kinds of converging lines in the retinal image as indicative of *parallel* edges in the environment. But in parallel projection, converging lines on the page can only indicate *converging* edges in the scene. Applying normal visual perception to such cases yields incorrect interpretations. Projection semantics captures the fact that depiction conforms with the general structure of vision, but unlike perceptual theories, allows that depiction also departs from vision in myriad ways.

Perceptual theorists may wish to speculate that the visual system is in fact flexible enough to incorporate non-optical systems of depiction in its production of perceptual representations. But then it would be relatively uninformative to say merely that pictorial content is determined by the exercise of vision. One would want to know, for different cases, *what kind* of vision is exercised, leading one back to an analysis in terms of specific methods of projection. In this case the perceptual theorist would be compelled to affirm the projection semantics anyway, understood as a semantic-level description of the varied capacities of the visual system.

I've canvassed three popular approaches to depiction. In each case, I've found the theoretical resources on offer to be either too imprecise or too rigid to usefully address the problem of pictorial space across a range of familiar systems of depiction. Although projection semantics only offers a partial account of pictorial representation, it has the virtue of accurately describing constraints on pictorial space for a wide variety of systems.

7 Beyond pure Projection

The primary target of this essay has been the problem of pictorial space for a central class of systems of depiction which I will call the **purely projective** systems. They all involve pictures which have a characteristically "mechanical" or "optical" look and feel, including nearly all forms of photography, "realistic" styles of drawing and painting, many of the drawing systems used in medical, engineering, architectural, and scientific contexts, as well as a broad range of projection-based maps. But pure systems do not exhaust the phenomenon of depiction, and in this section I turn my attention to the remaining **impurely projective** systems, in particular those which rely on stylization. I'll argue that the problem of pictorial space arises likewise for impurely projective systems, and that the same kinds of factors favor a semantics based on projection here. Yet, as we'll see, impure systems are incompatible with the present strict formulation of the Projection Principle in matters of spatial detail. Instead, my proposal is that impure systems of depiction should be explained by an extension of projection semantics, which applies more loosely to global spatial organization.

Purely projective systems of depiction are those whose characteristic methods of projection can be defined as functions of *only the local spatio-chromatic features* of the scene projected. Such methods depend only on **local** features of the scene in the sense that they can be defined entirely in terms of dependencies between very small regions of the scene and small regions of the marked picture plane. This means that the composition of the projected image can be worked out point by point (or small region by region) in relation to the scene. These projections depend only on **spatio-chromatic** features, like the presence of edges, surfaces, local textures, and reflectance properties. Linear perspective line drawing, for example, is purely projective because it can be defined as mapping the edge-features of point-sized regions in the scene to color-features of point-sized regions in the picture. Discussions of the projection-based approach to depiction in the philosophical literature have focussed almost exclusively on purely projective systems like this.

By contrast, impurely projective systems are those systems of depiction in which features of the picture plane cannot be derived from local spatio-chromatic properties of the scene alone. Instead, they are typically sensitive to high-level properties of the scene, such as whether it contains people, animals, or plants, or particular individuals. In such a system, how something is represented on the picture plane will depend on its high-level category. It turns out that there are many different kinds of impurely projective system. They include, by my count, systems of stylization, systems which involve more than one projective method, and systems which tolerate highly inconsistent or indeterminate geometry, like those found in some medieval depictions of architecture.[48] Though each of these cases is governed by its own set of rules, and deserves dedicated discussion, I'll focus here on stylization as representative of the broader class of impurely projective systems.

Before proceeding, stylization should be distinguished from **caricature**. Caricatures are pictures whose spatial content appears to be exaggerated in various ways, for various effects— to amuse, offend, facilitate recognition and so on Kris and Gombrich (1938), Tversky and Baratz (1985), and Rhodes et al. (1987). Though it may seem counter-intuitive, many forms of caricature are arguably purely projective, or nearly so. This is because they use distorted geometry in the picture plane to rep-

[48]By systems of "indeterminate geometry," I have in mind especially those which apply *aspects* of known pure projection systems, but inconsistently. This is characteristic of late medieval European depictions of space, prior to the codification of perspective during the Renaissance.

resent correspondingly distorted geometry in the pictorial space (Perkins 1975), just as a projection semantics would predict.[49] A gross caricature does not represent its subject as having, for example, normally shaped ears and eyes, but instead as having abnormally shaped ears and eyes. As a consequence, nearly all caricatures are deliberately inaccurate; indeed, caricatures seem to stand apart from other forms of purely projective depiction at the level communicative act, since we are not supposed to *believe* the contents of caricatures, but instead to treat them as representing reality in a distorted manner (Uidhir 2013).

Stylization arises when artists follow general prescriptions for accurately depicting different types of objects. Following Gombrich (1960, ch. 5) I call such prescriptions **schemas**. Schemas dictate different "ways" of drawing people, cats, trees, and so on. Such schemas are precisely what children learn when they learn, for example, how to draw a house or a tree. Rather than making their drawings conform principally to the way a given object is in fact shaped— as with pure projection— artists employing stylization instead make their drawing conform with an antecedently established schema. Far from a specialized case, stylization dominates art history. Virtually all pictures created before the Italian Renaissance contain some degree of stylization, and a large share of picture-production since continues this tradition. Here are two forms of stylization, one from a contemporary crossing sign, the other from a tomb in the Theban Necropolis, dating from the 15th century BCE.

N

O

Whatever the precise spatial content of Picture N, it is clear that it depicts a normally-shaped person walking on a path or crosswalk.[50] The fact that the head of the stick figure is drawn with a large black circle does not imply that, in the content, there is a large sphere floating above a stick-like form. Rather, this apparent violation of

[49] That said, many forms of caricature do blend visual exaggeration with stylization, in which case the comments on stylization below would apply.

[50] A different diagnosis holds that such images are "distortions" of their subjects, which is to say, they are systematically and grossly inaccurate (cf. Voltolini 2015, 63-64). This would bring the analysis of stylization in line with my analysis of caricature, and consequently obviate the theoretical challenge facing the projective analysis. Yet stylization seems to me too ubiquitous, and its uses too often sincere and factual, to dismiss with a widespread error-theory. Stylized images are used to record fact in a way that caricatures are not. For better or worse, I think that projection semantics is stuck with the problem of stylization.

projection is the result of the application of a schema for drawing humans. Similarly, Picture O depicts a group of men in a funeral procession. The fact that, in the picture, we see the frontal plane of their chests, but the sides of their feet, does not imply that their bodies are impossibly twisted around (Hagen 1986, 169-170). Instead, they are again produced from an antecedent prescription for drawing people, in a manner to some degree independent of their actual shape. In general, to the extent that stylized images employ such schemas, they may depart from pure projection.[51]

Nevertheless, I believe that stylization is grounded in projection in some definite but less strict sense. The evidence is much the same as in the purely projective case. Stylized images clearly express pictorial spaces: they locate objects and properties at different directions and depths relative to an implied viewpoint. These pictorial spaces vary by system of depiction in familiar ways. Some stylized images clearly express spaces which conform approximately with perspective projection, as in the illustration from the 1943 Belgian comic *The Secret of the Unicorn* at left; while others conform with parallel-projection, as in the panel from the 1756 Japanese wood-block illustration of the classic *Tales of Ise* at right. This is just the now familiar problem of pictorial space for the case of stylized pictures. That both images are stylizations can be seen from the stereotyped manner in which people's faces, among other elements, are depicted in each case.

Something must explain the mapping from individual stylized pictures to the pictorial spaces they express. Given the apparent variation between perspective-like and parallel-like spaces among stylized images, we should expect an answer similar to the one deduced for purely projective systems. That is, something very much like a projection principle which allows for variation by methods of projection must explain the variation in pictorial space in these cases. Of course, the principle in ques-

[51] That said, the schemas adopted in stylized depiction are not arbitrary. Instead, they are loosely derived from projections from stereotyped viewpoints. The bubble-figure in the crosswalk sign illustrates this principle, for the overall organization of legs, arms, torso, and head correspond loosely to the positions those objects would have in a pure projection from a side view. (A projection from above or below would have a rather different layout.) In general, in a stylized system, the schema for a type of object X is derived from the pure projection(s) of a standard instance of X from some canonical viewpoint (or viewpoints). See Hagen (1986, ch. 6) and Willats (1997, 212-214) for discussion of canonical viewpoints.

tion cannot require *local* projective coherence between the picture and the space it expresses; such local coherence is broken by the use of schemas. Instead the principle must impose only a looser *global* projective coherence that constrains the depicted relationships between objects, even if it doesn't govern the depiction of objects themselves. I leave it to future work to spell out this idea in detail.

Meanwhile, the competing resemblance and perceptual theories do not seem to be able to account for even the global features of pictorial space in stylization, and for familiar reasons. Capturing the geometrical differences between curvilinear, parallel, and perspective spaces is, as I have argued, a central faltering point for both resemblance and perceptual theories of depiction. We should expect the same dialectic to play out for stylized images, again favoring some form of projection semantics. Thus contextual resemblance will still face the challenge of stylized transformative systems (like curvilinear perspective or shifted color). And perceptual resemblance accounts, as well as perceptual accounts, will still face the challenge of adequately capturing the content of stylized parallel projections. In general, if a theory cannot account for aspects of a purely projective system, it will also fail for its stylized cousin. The apparent explanatory value of a global projective constraint, and the apparent challenges facing alternative accounts, give us reason to expect that the semantics of stylized systems will be constrained by projection at least at the level of global spatial organization.

It remains to be seen whether the proposal outlined here will extend to other types of impurely projective system, like those which combine multiple methods of projection, or those with inconsistent geometry. But here again the ground level facts are much the same: such pictures express pictorial spaces, and something must explain the attribution of spaces to pictures. A global projection requirement of some kind would satisfy this explanatory demand and unify the treatment of pure and impure systems.

8 Conclusion

The argument of this essay can be understood as an extended inference to the best explanation. That which is to be explained is the problem of pictorial space: how pictures in different systems express the variety of pictorial spaces. The hypothesis offered to address this problem is projection semantics. I've shown how this explanation works, and I've provided an account of the relationship between vision and projection that helps explain why humans would adopt such a semantics in the first place. Finally, I've argued that alternative theories of depiction provide less successful accounts of the same data. For purely projective systems, projection semantics provides the best account of the directional layout of basic features. Since directional structure is the basis of pictorial space, and pictorial space is the basis for high-level pictorial content, projection semantics is the natural starting point for any further pictorial semantics.

Looking beyond pure systems, I've shown how the same pattern of explanatory motivations and challenges arise for systems of stylization. We have reason to think that a global projective constraint is playing a role here as well. These observations support the general hypothesis that all forms of pictorial representation, both pure and impure, have semantics which are grounded in a global projection constraint. This would explain why all forms of pictorial content seem to take the form of viewpoint centered spaces, and why the structure of these spaces vary in familiar ways, by known methods of projection. It would imply that the projective structuring of space is one of the defining characteristics of pictorial representation, distinguishing it from both diagrammatic and linguistic representation. My conjecture, then, is that a semantics based in projection is a central mark of pictorial representation.

Acknowledgments This paper has grown out of years of discussion with dozens of friends, colleagues, teachers, students, and audiences. I owe far too many personal and intellectual debts to enumerate here. Parts of this research were supported by NSF IGERT DGE 0549115.

Appendix: Bare bones content and the Projection Principle

In this appendix I argue that, in the context of Kulvicki's (2006, ch. 3) claims about pictorial structure, the constraint he offers on pictorial content is significantly weaker than the Projection Principle.

To begin, I restate the Projection Principle within the framework of possible-world semantics. Where S is a system of depiction and c is a context, $[\![\cdot]\!]_{S,c}$ is an interpretation function mapping pictures to sets of scenes, understood as $\langle world, viewpoint\rangle$ pairs. For a system S, $proj_S(\cdot)$ is the characteristic projection function of S, mapping scenes to pictures. Then according to the Projection Principle, given a system S and context c, for any picture P in S:

$$\forall w \forall v: \text{ if } \langle w, v\rangle \in [\![P]\!]_{S,c} \text{ then } proj_S(w, v) = P \tag{1}$$

Next, Kulvicki posits a basic layer of pictorial content, called **bare bones content**, which in turn is defined in terms of projection: "whatever scenes could have resulted in a particular [linear perspective picture] via a perspective projection count as parts of the [bare bones] content of the picture Kulvicki (2006, 59)." This formulation, in fact, expresses only the sufficiency of projection for bare bones content, but the necessity clause is clearly implied as well. So we can think of the bare bones content of a picture as the set of all scenes which project to the picture, relative to a sytem. Letting $B_S(\phi)$ be the bare bones content of ϕ, relative to system S, we can render Kulvicki's definition within a scene-based semantics, as follows:

$$B_S(P) = \{\langle w, v\rangle \mid proj_S(w, v) = P\} \tag{2}$$

Kulvicki further holds that bare bones content constrains a picture's overall content. That is:

$$[\![P]\!]_{S,c} \subseteq B_S(P) \tag{3}$$

Put together, these entail:

$$[\![P]\!]_{S,c} \subseteq \{\langle w, v\rangle \mid proj_S(w, v) = P\} \tag{4}$$

which is equivalent to the Projection Principle.

However, Kulvicki is motivated to preserve a property of systems he terms **transparency** (Kulvicki 2006, ch. 3). Intuitively, transparency is preserved when a picture *of* a picture is (syntactically) the same as the original. Formally: a system S is transparent iff, for any representations R and R' from S, if R' satisfies the content of R, then R and R' are of the same syntactic type (p. 53). Kulvicki seeks to establish that core systems, including linear perspective, satisfy this constraint. Yet he acknowledges that the relevant level of content is not plausibly the full, "fleshed-out" content of a picture (p. 52). But it cannot even be its bare bones content, as defined above, because if P' is projected from P relative to an oblique viewpoint, P' will satisfy P's bare bones content, but the two have manifestly different metrical structures (pp. 53-55).

To solve this problem, and save transparency, Kulvicki proposes to revise the operative notion of pictorial syntax. Rather than defining pictorial syntax metrically, as is standard, he holds that pictorial syntax consists only of projective invariants. Since pictorial syntax is the only feature of a picture which determines bare bones content, it now follows that all pictures which share projective invariants have the same bare bones content (Kulvicki 2006, 57).

To compare the resulting idea with the Projection Principle, I'll translate Kulvicki's proposal back into the framework of this essay in which pictures are treated metrically. Kulvicki's (final notion of) bare bones content can be defined as the set of all scenes which project to something which shares projective invariants with the target picture. Since projections of a picture always share projective invariants with it, I can restate the final definition of bare bones content, now B^*, as follows. Here I'll allow that a picture P defines a world w_P which contains only that picture.

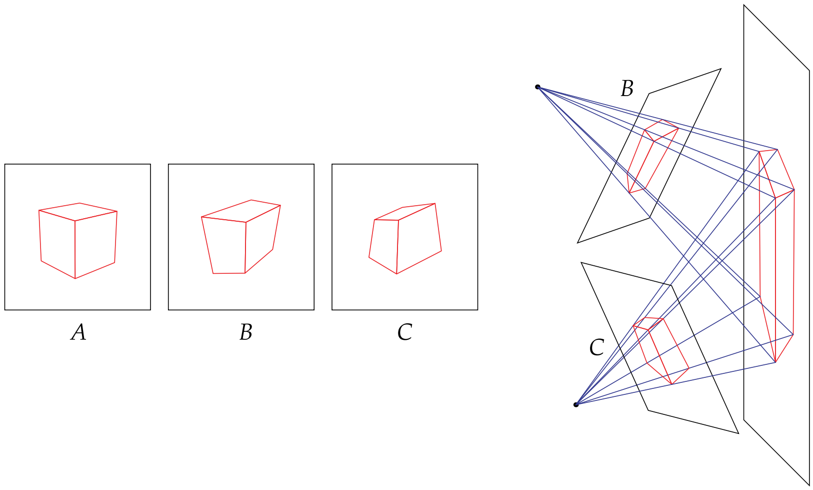
$$B^*_S(P) = \{\langle w, v\rangle \mid \exists v' : proj_S(w, v) = proj_S(w_P, v')\} \tag{5}$$

The resulting constraint on content can now be set side by side the Projection Principle:

$$\text{if } \langle w, v\rangle \in [\![P]\!]_{S,c} \text{ then } \exists v' : proj_S(w, v) = proj_S(w_P, v') \qquad \text{(Kulvicki's Constraint)} \tag{6}$$

$$\text{if } \langle w, v\rangle \in [\![P]\!]_{S,c} \text{ then } proj_S(w, v) = P \qquad \text{(Projection Principle)} \tag{7}$$

The difference here is significant. Consider the diagram below in which pictures B and C are projected from A, thus share projective invariants with it:



Intuitively, A, B, and C have *different* spatial content. The Projection Principle delivers this result: since each image can be projected from scenes which the others could not, they do not have the same skeletal content, hence they do not have the same content. The content of A, for example, is compatible with scenes containing a cube, but this is not the case for the contents of B or C.

But Kulvicki's Constraint cannot capture this claim. Let L be the system of perspective projection. Perspective projection from a plane is reversible, modulo issues of scale; so if there is some v such that $proj_L(w_P, v) = P'$, then there is some v' such that $proj_L(w_{P'}, v') = P$. As a consequence, Kulvicki's Constraint puts the same extensional requirement on each of A, B, and C. This is to be expected, given the background theory: if all three have the same pictorial structure, and structure determines bare bones content, then Kulvicki's Constraint can't distinguish between them. Thus the Projection Principle issues in a more determinate constraint that Kulvicki's Constraint.

While the semantic differences between A, B, and C are in principle compatible with Kulvicki's view at the level of what he calls "fleshed-out content," only the Projection Principle actually delivers the relevant result. Since neither Kulvicki nor I have offered a detailed account of how fleshed-out content is determined, the Projection Principle goes farther, given the theoretical resources available, towards explaining these differences in a systematic manner.

References

Abell, C. 2009. Canny Resemblance. *Philos Rev* 118.2: 183–223.

Abusch, D. 2015. Possible worlds semantics for pictures. In: Blackwell Companion to Semantics, eds. L Mathewson, and et al. New York: Wiley. Forthcoming.

Beck, J. 2015. Analogue magnitude representations:Aphilosophical introduction. *Brit J Philos Sci* 66.4:829–855.

Blumson, B. 2009. Pictures, Perspective and Possibility. Philos Stud 149.2:135–151.

Blumson, B. 2014. *Resemblance and Representation: An Essay in the Philosophy of Pictures*. Open Book Publishers.

Budd, M. 1996. How pictures look. In: Values of art: Pictures, poetry, and music: Blackwell.
Casati, R, and V Giardino. 2013. Public Representation and Indeterminicies of Perspectival Content. In: Enacting Images, ed. Z Kondor, 111–126: Herbert von Halem Verlag.
Chomsky, N. 1965. *Aspects of the Theory of Syntax*. The MIT press.
Cutting, JE. 2003. Reconceiving perceptual space. In:.
DeCarlo, D, and et al. 2003. Suggestive Contours for Conveying Shape. *ACM Trans Graph (TOG)* 22.3:848–855.
DeLoache, JS, SL Pierroutsakos, and DH Uttal. 2003. The origins of pictorial competence. *Curr Direct Psychol Sci* 12.4: 114–118.
Dretske, F. 1981. Knowledge and the Flow of Information, MIT Press, Cambridge.
Dubery, F, and J Willats. 1972. *Drawing systems*. Studio Vista, London.
Durand, F. 2002. An invitation to discuss computer depiction. In: In: Proceedings of the 2nd international symposium on Non–photorealistic animation and rendering, 111–124: ACM.
Firestone, C, and BJ Scholl. 2016. Cognition does not affect perception: Evaluating the evidence for "top–down" effects. *Behav Brain Sci* 39.
French, S. 2003. A Model–Theoretic Account of Representation (or, I Don't Know Much about Art – but I Know it Involves Isomorphism). *Philos Sci* 70.5: 1472–1483.
Frisby, JP, and JV Stone. 2010. *Seeing: The computational approach to biological vision*. The MIT Press.
Giardino, V, and G Greenberg. 2015. Varieties of Iconicity. *Rev Phil Psychol* 6.1: 1–25.
Gombrich, EH. 1960. Art and Illusion: A Study in the Psychology of Pictorial Representation. Phaidon, London.
Goodman, N. 1968. *Languages of Art: An Approach to a Theory of Symbols*. Bobbs–Merrill, Indianapolis.
Greenberg, G. 2011. *The Semiotic Spectrum*. PhD thesis. Rutgers University.
Greenberg, G. 2013. Beyond Resemblance. *Philos Rev* 122.2: 215–287.
Greenberg, G. 2018. Content and Target in Pictorial Representation. *Ergo* 5.23.
Greenberg, G. 2019. Tagging: Semantics at the Iconic/Symbolic Interface. In: Proceedings of the.
Greenberg, G. 2020. The Structure of Visual Content. Manuscript.
Hagen, MA. 1986. *Varieties of Realism: Geometries of Representational Art*. Cambridge University Press, Cambridge.
Hammer, E, and S Shin. 1998. Eulers visual logic. *History Philos Log* 19.1: 1–29. issn: 0144–5340.
Hansen, R. 1973. This CurvingWorld: Hyperbolic Linear Perspective. *J Aesthet Art Criticism* 32.2: 147–161.
Hopkins, R. 1998. *Picture, Image and Experience: A Philosophical Inquiry*. Cambridge University Press, Cambridge.
Hopkins, R. 2003. Perspective, convention and compromise. In: Looking Into Pictures: An Interdisciplinary Approach to Pictorial Space, eds. M Atherton, R Schwartz, and H Hecht: MIT Press.
Hopkins, R. 2006. The speaking image: Visual communication and the nature of depiction. In: Contemporary Debates in Aesthetics and the Philosophy of Art, ed. M Kieran: Blackwell.
Howell, R. 1974. The logical structure of pictorial representation. *Theoria* 40.2: 76–109.
Hyman, J. 2006. *The Objective Eye: Color, Form, and Reality in the Theory of Art*. University of Chicago Press.
Hyman, J. 2012. Depiction. *R Instit Philos Suppl* 71: 129–150.
Ittelson, WH. 1996. Visual perception of markings. *Psychon Bullet Rev* 3.2: 171–187.
Johnson, GM. 2020. The structure of bias. In: Mind.
Kennedy, JM. 1974. *A psychology of picture perception*. Jossey–Bass Publishers, London.
Koenderink, J, and A van Doorn. 2003. Pictorial space. In: eds. H Hecht, R Schwartz, and M Atherton. Cambridge: MIT Press.
Korsmeyer, C. 1985. Pictorial assertion. *J Aesthet Art Criticism* 43.3: 257–265.
Kris, E, and EH Gombrich. 1938. The Principles of Caricature. *Brit J Med Psychol* 17.3–4: 319–342.
Kulvicki, J. 2006. *On Images: Their Structure and Content*. Clarendon, Oxford.
Lewis, D. 1975. Languages and Language. In: Language, Mind, and Knowledge. Ed. by K. Gunderson. Vol. 7. Minnesota Studies in the Philosophy of Science. University of Minnesota Press, 3–35.
Leyton, M. 1992. *Symmetry, causality, mind*. The MIT Press.
Lopes, DM. 1996. *Understanding Pictures*. Oxford: Oxford University Press.
Lopes, DM. 2006. The domain of depiction. In: Contemporary Debates in Aesthetics and the Philosophy of Art, ed. M Kieran: Blackwell.
Marr, D. 1982. *Vision*. New York: Henry Holt and Co.

Neander, K. 1987. Pictorial Representation: A Matter of Resemblance. *Brit J Aesthet* 27.3: 213–226.
Newall, M. 2011. *What is a Picture?: Depiction, Realism, Abstraction*. Palgrave Macmillan.
Palmer, SE. 1999. *Vision Science: Photons to Phenomenology*. The MIT Press.
Peacocke, C. 1987. Depiction. *Philos Rev* 96.3: 383–410.
Peacocke, C. 1992. *A Study of Concepts*. The MIT Press.
Perkins, D. 1975. A definition of caricature and caricature and recognition. *Stud Vis Commun* 2.1: 1–24.
Pirenne, M. 1970. *Optics, painting & photography*. Cambridge University Press.
Pylyshyn, Z. 1999. Is vision continuous with cognition? The case for cognitive impenetrability of visual perception. *Behav Brain Sci* 22.3: 341–365.
Rhodes, G, S Brennan, and S Carey. 1987. Identification and ratings of caricatures: Implications for mental representations of faces. *Cogn Psychol* 19.4: 473–497.
Robinson, AH, and et al. 1995. *Elements of Cartography*, 6th Edn. Wiley, New York.
Rock, I. 1983. *The Logic of Perception*. MIT Press.
Rogers, B, and C Rogers. 2009. Visual Globes, Celestial Spheres, and the Perception of Straight and Parallel Lines. *Perception* 38.9: 1295–1312.
Rogers, S. 1995. Perceiving pictorial space. *Percept Space Motion* 5.
Rogers, S, and et al. 2003. Truth and meaning in pictorial space. In: Looking into pictures: An Interdisciplinary Approach to Pictorial Space: 301–320.
Ross, J. 1997. *The Semantics of Media*. Kluwer Academic Publishers.
Schier, F. 1986. *Deeper Into Pictures: An Essay on Pictorial Representation*. Cambridge University Press.
Scruton, R. 1981. Photography and representation. *Crit Inquiry*: 577–603.
Sedgwick, H. 1980. The Geometry of Spatial Layout in Pictorial Represenation., (Vol. 1. In: In: The Perception of Pictures, 33–90. London: Academic Press.
Shin, S. 1994. *The Logical Status of Diagrams*. Cambridge University Press.
Sober, E. 1976. Mental representations. *Synthese* 33.2: 101–148.
Treisman, A. 1988. Features and objects: The fourteenth Bartlett memorial lecture. *Quart J Exper Psychol* 40.2: 201–237.
Tversky, B, and D. Baratz. 1985. Memory for faces: Are caricatures better than photographs? *Memory Cogn* 13.1: 45–49.
Tye, M. 2000. *The imagery debate*. The MIT Press.
Uidhir, CM. 2013. The Epistemic Misuse & Abuse of Pictorial Caricature. *Amer Philoso Quart* 50.2: 137–152.
Vishwanath, D, AR Girshick, and MS Banks. 2005. Why pictures look right when viewed from the wrong place. *Nat Neurosci* 8.10: 1401–1410.
Voltolini, A. 2015. *A syncretistic theory of depiction*. Springer.
Walton, K. 1973. Pictures and Make–Believe. *Philos Rev* 82.3: 283–319.
White, J. 1957. *The birth and rebirth of pictorial space*. Cambridge, MA.
Willats, J. 1997. *Art and Representation: New Principles in the Analysis of Pictures*. Princeton University Press, Princeton.
Wollheim, R. 1987. *Painting as an Art*. Thames and Hudson.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.