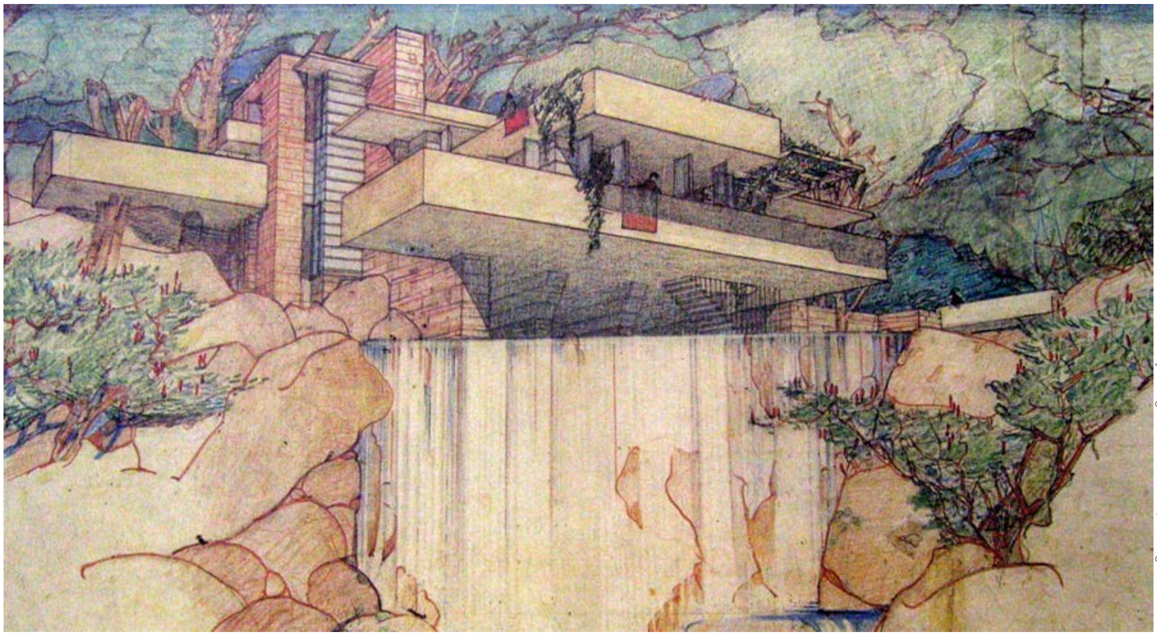
Pictorial Semantics

Gabriel Greenberg 1.12.12 Department of Philosophy, University of California Los Angeles



Frank Lloyd Wright and Jack Howe, "Fallingwater" 1953

Because humans cannot know one another's minds directly, every form of communication is a solution to the same basic problem: how can privately held information be made publicly accessible through manipulations of the physical environment? Language is by far the best studied response to this challenge. But there is a diversity of non-linguistic strategies for representation with external signs, from facial expressions and fog horns to chronological graphs. Among these alternatives, the class of *pictorial* representations dominate practical communication— witness the proliferation of maps, road signs, newspaper photographs, scientific illustrations, television shows, architectural drawings, and even the fleeting imagery of manual gesture.[1] In such cases, what bridges the gap between the physical media of pictures and the information they manage to convey? In this paper I offer a partial answer to this question.

The general thesis of this paper is that pictures are associated with their informational content by systematic rules selected by social convention— not, as one might have thought, by a purely biological reflex of vision, nor by open-ended rational speculation. In this respect, the interpretation of pictures is importantly similar to the interpretation of sentences in a language. I will call this the *Pictorial Semantics Hypothesis*.[2] In addition to this general claim, I will defend a specific and

Acknowledgements. This paper has benefited from the generous conversation and comments of many friends, colleagues, and teachers: Tendayi Achiume, Luvell Anderson, Josh Armstrong, Sam Cumming, Pavel Davydov, Doug DeCarlo, Richard Dub, Thony Gillies, Edinah Gnang, Chip Greenberg, Rohit Gupta, Chris Hom, Alex Jackson, Michael Johnson, Philip Kremer, Ernie Lepore, Karen Lewis, Dom Lopes, Ricardo Mena, Zachary Miller, Lisa Miracchi, Alex Morgan, Sarah Murrey, Carlotta Pavese, Kevin Sanik, Will Starr, participants in the NSF-IGERT Perceptual Science Program at Rutgers University, and audiences at the 2010 Eastern Division of the American Society for Aesthetics, the 2010 CEU Summer School on Content and Context in Budapest, and the University of Johannesburg. Special thanks are due to my advisors for their expert assistance and enduring patience: Jeff King, Jason Stanley, Peter Kivy, and Matthew Stone. This research was supported in part by NSF IGERT DGE 0549115, under the guidance of Matthew Stone.

[1]See McCloud (1993), Tversky (2000), and Lascarides and Stone (2009a, 2009b) for a wide range of examples.

[2]I shall simply assume that natural languages have semantics, but even this is still contested. Stanley (2000) and King

largely novel proposal about what kind of semantics pictures actually have. According to this idea, for a picture to accurately depict a scene is for there to be a projective mapping from the three-dimensional features of the scene to the two-dimensional features of the picture plane. Recovering a picture's content from its surface features involves working out the kind of scene from which the picture could have been projected. In this respect, the strategy used to encode information in pictures is decidedly unlike that employed by languages. I will call this the *Geometrical Projection Theory*.[3]

If correct, these conclusions have import for the study of a variety representational kinds beyond the pictorial. By providing a precise formal framework in which geometrical structures are associated with correspondingly geometrical content, they show that rigorous semantic analysis can assume forms quite unlike those most familiar today. Languages constitute just one class among a broad spectrum of representational systems, all of which deserve to be studied under a general science of semantics. In addition, this framework may serve as a model for the philosophical analysis of visual perception. Just as the study of public languages has informed our understanding of private cognition, the present account of public images may help to guide the philosophic theory of perceptual representation generally.[4]

The argument of the paper is unfolded in three parts, organized from the most general statement to the most specific. Part 1 presents and defends the Pictorial Semantics Hypothesis, the view that pictures are associated with their content by conventional systems of interpretation. Part 2 introduces the Geometrical Projection Theory, a particular account of the semantics for pictures, and contrasts it with the main alternative proposals from the Philosophy of Art. Part 3 develops this theory in precise detail for the particular system of perspective line drawing. The last section is a conclusion. In the Appendix, the proposed semantics is formalized and shown to be fully compositional.

1 The Pictorial Semantics Hypothesis

The use of languages in human society is universal. Languages, manifested variously in speech, writing, and bodily signs, enable the fast and reliable communication of complex information between individuals. It is widely believed that languages are based on complex and systematic mech-

---

and Stanley (2005) survey such anti-semantic accounts of language and argue against them. The present paper takes the same basic considerations that motivate King and Stanley and applies them to the case of pictures.

[3] There are, to my knowledge, only a very few explicit proposals for semantic accounts of pictorial representation. They include Malinas (1991) and Mackworth and Reiter (1989). Setting aside its affiliation with semantics, the novelty of the Geometrical Projection Theory depends on the comparison class. In projective geometry and perceptual science something like this account is simply *assumed* (see, e.g. Hagen 1986, Willats 1997, Durand 2002). But such inquiry tends to be inexplicit about the notion of depiction under analysis. On the other hand, in philosophical discussion, where the subject is more clearly set out, the theory has little precedent. The most similar account is that of Hyman (2006, ch. 6). The present theory generalizes, precisifies, and diverges from Hyman's proposal in a number of ways; see footnote **??** for discussion. A related account, which nevertheless differs fundamentally from the present theory is that of Kulvicki (2006, ch. 3-4); see footnotes **??** and **??**.

[4] There are other potential applications as well. A precise semantics of pictures would be illuminating for the study of imagistic mental representation— whether or not the suggestive analogies between these domains are taken literally. (Block 1982, 1-16; Pylyshyn 2006, ch. 7) In the philosophy of science, debate has arisen recently about the nature of scientific models and scientific representation in general; here pictures are frequently held up as a paradigm case (see, e.g. French 2003; Giere 2004, §4; Cohen and Callender 2006, 11). Even in metaphysics a pictorial semantics might elucidate the metaphysical account of possible worlds known as "pictorial ersatzism." (Lewis 1986, §3.3)

anisms which regulate the production, structure, and meaning of linguistic signals. The study of SEMANTICS, in particular, hypothesizes the existence of tacitly understood rules which govern meaning— rules which map linguistic signs to representational content. Such theories contribute to an explanation of linguistic communication by describing mechanisms by which language users employ shared conventions to encode and decode information in linguistic signs. Since the 1960's and 70's the mathematical study of semantics has flourished, producing detailed predictions about an ever-widening range of linguistic data. Now one of the central pillars of modern linguistics, semantic theory has proven to be a progressive and exact research program.

Of course, there are other, non-linguistic forms of representation as well, most notably, representation by pictorial images. The use of pictorial representations as tools of public expression is ancient in human history, and thrives without special training or tools in the form of iconic gesture. In modern industrial society, pictures are ubiquitous, used to efficiently encode and transmit vast quantities of information— exemplified by maps, road signs, text book illustrations, architectural drawings, television broadcasts, and so on. And the roots of pictorial representation run deep: the spatial organization of the human visual cortex strongly suggests that picture-like representation is one the basic strategies for information management implemented by the brain, particularly in low-level perceptual processing.[5]

The first major thesis of the present paper is that we should posit a semantics for pictures, just as we have posited a semantics for natural language. The proposal is not, implausibly, that pictorial and linguistic representation are governed by the same *kind* of semantics— only that both are governed by semantics.[6] More explicitly:

(1) ***Pictorial Semantics Hypothesis***
Pictures are associated with their content by systematic and conventional interpretive rules.

This Hypothesis consists of three key claims, each of which will be defended separately in what follows: (i) that pictures have representational content; (ii) that the association between pictures and their content is systematic; (iii) that the systematic association between pictures and their content is conventional.

The Pictorial Semantics Hypothesis has few allies. Both philosophers of language *and* philosophers of depiction have doubted that non-linguistic representation is amenable to the same kind of semantic analysis that characterizes language. As a consequence, representation by pictorial images is widely dismissed as a legitimate subject of systematic study. Those who are moved by such skepticism must admit that pictures often play a pivotal role in successful communication, but they insist that pictorial communication is wholly accounted for by some admixture of facts about human visual perception, general principles of rationality, and heterogeneous world knowledge. In short, they believe there is a *pragmatics* of pictorial representation, but not a semantics. Thus Walton (1990, 351) writes that language "is to be defined in semantic and/or syntactic terms," but "depiction is a pragmatic notion." And Stanley (2000, p. 396) concludes, "we should therefore be suspicious of attempts to forge philosophically significant analogies between the different

[5] See Pylyshyn (2006) for extended discussion. It remains a controversial question whether there are also picture-like representations in higher cognition.

[6] On the general conception of semantics see Tarski (1944, 345) and Haugeland (1991, 62).

processes underlying the interpretation of linguistic and non-linguistic acts."[7]

What accounts for the widespread skepticism about pictorial semantics— given the widespread *acceptance* of linguistic semantics? One reason, I believe, is that the evidence for pictorial semantics has not been fully appreciated; one of the tasks of the first half of this paper is to review this evidence. There are also a number of substantive conceptual and empirical worries one might have about a pictorial semantics; another task of the present discussion is to address these concerns. There are still other sources of doubt which stem from basic misunderstandings about the project of pictorial semantics. In the remainder of this section, I address three such preliminary issues, with an eye towards clarifying the core aims of the present investigation.

First and foremost is the simple conflation of *pictures* and *visual art*. The interpretation of works of art is by nature unbounded, in part because artistic meaning is specifically tuned to expressive and metaphorical significance beyond literal content, in part because such meaning is often a product of intentionally violating the very interpretive norms that govern ordinary communication. This makes artistic meaning an admittedly unlikely candidate for systematic semantic analysis. But just as linguists do not take poetry as their primary source of data, there is no reason that students of pictorial representation should take visual art works as their point of departure. Instead, we should only expect robust semantic regularities to emerge from the use of pictures deployed for practical information exchange, under conditions in which efficiency and fidelity are at a premium. Examples include road signs, maps, architectural and engineering drawings, textbook illustrations, police sketches, and so on. In this paper I will confine my attention exclusively to pictures which are intended to answer to such standards; works of art are not my subject matter. (Still, I will use the term "artists" to refer generically to the creators of pictures.)

A second reason for doubt is the assumption that, as a matter of definition, semantics has the structure of linguistic semantics. Yet it is correctly observed that the mechanisms underlying pictorial representation are not those of linguistic representation. Languages encode representational content, in the first instance, via arbitrary associations of individual words with meanings; the meanings of sentences are derived from word meanings by rules of composition and context-sensitivity. But this linguistic model is inappropriate for pictorial representation for many reasons. The most fundamental is that successful pictorial representation does not seem to be arbitrary at all. The relationship between a drawing or photograph of, say, a tree, and the tree itself is one of intimate geometrical correspondence, nothing like the stipulative association between the word "tree" and the property of being a tree. So, the objection goes, since semantics is defined in terms of the the atomic and arbitrary structure of language, and since pictorial representation resists such definition, semantics is not for pictures.

But this argument assumes an obtuse vision of pictorial semantics. Insofar as the idea of a pictorial semantics is plausible, it is not the idea of a semantics whose architecture is just like that of language. Rather, by abstracting away from the atomic and arbitrary basis of linguistic representa-

[7] And Borg (2006, 261) announces "fundamental differences between communicative acts in general and linguistic acts in particular. Linguistic acts, uniquely in this area, have a crystallized component to their meaning, an element which they carry with them across all contexts and which may be accessed by a competent language user even if she has no access at all to the speaker's original intentions... Yet these genuinely code-like qualities seem very different to the properties of other communicative acts, which depend on context in a far more constitutive way."

tion, we can isolate a more general concept of semantics. The Pictorial Semantics Hypothesis posits a semantics for pictures only in this ecumenical sense— a systematic and conventional mapping from signs to content.

A third and final concern is that none of the extant theories of depiction could plausibly form the basis of a systematic pictorial semantics. For many theorists, this result is quite deliberate. For example, the *Symbol Theory* of depiction introduced by Nelson Goodman in his 1968 *Languages of Art* was articulated with groundbreaking rigor, but it rejected the view that depiction was in any substantive sense systematic. Goodman held the implausible view that pictures are associated with their content only by arbitrary stipulations, much like the relation between words and their denotations. A much more promising tack is pursued by *Perceptual Theories*, which define the content of a picture in terms of the cognitive and perceptual effects it has on viewers. Yet in general, such theories purposefully eschew the systematicity of semantic analysis, holding instead that an exact definition of depiction ultimately rests in large part on the gnarly mechanics of biological vision. The only theory of depiction which immediately lends itself to a semantics is the orthodox *Resemblance Theory*, according to which a picture represents a scene in virtue of being *similar* to it in certain relevant respects. But such accounts are generally products of philosophical dialectic, not quantitative empirical inquiry. And they have been the subject of numerous and vigorous objections over the last fifty years. In Greenberg (2010), I add my voice to this dissenting chorus, arguing that the the very structure of such similarity-based views forces them into false predictions for a broad range of cases. But if none of the extant philosophical theories of depiction could form a suitable basis for a pictorial semantics, how could the Pictorial Semantics Hypothesis possibly be right?

To vindicate the Pictorial Semantics Hypothesis, we will need to identify some other, novel account of depiction to answer to the requirements of a pictorial semantics. And it is this mandate which the second half of this paper aims to fulfill. Inspired by the study of images in projective geometry and recent work in computational perceptual science, the Geometrical Projection Theory which I'll defend diverges from its predecessors both in the substance of its analysis and the exactness of its expression. According to this theory, a picture depicts a scene just in case the picture can be derived from the scene by general but precise rules of geometrical projection. Thus pictorial representation is grounded neither in arbitrary meaning associations, subtle resemblances, nor the perceptual responses of viewers. The result is an account that boasts a rigor, simplicity, and empirical validity which has eluded other proposals; it is an appropriate foundation for a semantics of pictorial representation.

I hope to have clarified the rudimentary commitments and ambitions of the Pictorial Semantics Hypothesis. In the remainder of this first half of the paper I will elaborate and defend each of its three key components: (i) the presupposition that pictures have representation content; (ii) the posit of systematic mappings from pictures to their content; and (iii) the claim that these systematic mappings are established in human communication by social convention.

1.1 Pictorial content

To begin, a terminological note: a PICTURE, as I use the term, is a *type* of pictorial sign, not a concrete token. I intend to include typical cases of photography and so-called "representational" (as opposed to abstract) painting and drawing. But I will not consider composite images, such as comic books, or images which integrate symbolic elements like textual labels or color-coding. For purposes of brevity, I shall also ignore such abstract representations as cartesian graphs, pie charts, and Venn diagrams— as well as alternative media such as cartoons, film, relief, scale models, music, and audio recording.[8]

The Pictorial Semantics Hypothesis presupposes that pictures have REPRESENTATIONAL CONTENT. In this section I will first elaborate and then defend this position.[9] At minimum, by claiming that pictures have representational content, I mean that pictures are representations which specify a way reality could be, with respect to objects and properties that can be physically remote from the picture itself. The content of a picture is identical to the way reality could be, as specified by the picture.[10] The opposing view holds that pictures are ordinary, non-representational objects; their communicative power stems only from their causal effects on the minds of viewers.

The idea of pictorial content at work here is familiar enough. Of Frank Lloyd Wright and Jack Howe's 1935 architectural drawing with which we began, it is natural to say that it depicts a house, of a certain shape and color, in a certain relationship to its environment, perched over a waterfall, and so on. In this way, we commonly associate pictures— flat, marked surfaces— with specifications of reality as having specific, spatially extended features. That is, we naturally associate pictures with representational content.

Content in this sense is not an aesthetic or impressionistic property of pictures. Like sentences, images can reliably carry precise and useful informational content. For example, the exact content carried by the Frank Lloyd Wright drawing, about a hypothetical building plan, played a crucial part in the Kauffman family's eventual decision to build the house, at considerable cost.[11] In general, the use of pictures to communicate specific information in high-stakes contexts, such as those encountered by engineers, architects, and pilots, demonstrates the efficacy of pictorial content, and should quell the suspicion that the content of pictures is in any way less exact or stable than that of sentences.

Pictorial content must be distinguished from other semantic properties of pictures. First, the content of a picture is not its REFERENT. Whereas the content of a picture is the depictive information it carries, its referent is the scene which the picture is supposed to carry information *about*. While pictorial reference and content both answer to colloquial uses of the phrase "what is depicted", they come apart strikingly in cases of misrepresentation.[12]

[8] The discussion here can be recapitulated for nearly any of these media. See Kivy (1984, ch. 1-2) for a detailed parallel discussion of musical content.

[9] The general thrust of this section— that pictures have content much like sentences, and that this content can be defined in terms of correctness conditions— is clearly prefigured in Wittgenstein (1921/1997, §2.21-224).

[10] Since this characterization does not put direct constraints on the nature of content, it is compatible with the many divergent theories of content on offer in contemporary debates. See Siegel (2009, §3) and King (2007, ch. 1) for discussion.

[11] See Toker (2003, 195, 250-7).

[12] The distinction between pictorial reference and pictorial content is widely observed in the Philosophy of Art; see for example Beardsley (1958, 270-2), Goodman (1968, 27-31), and Lopes (1996, 151-2). Various ways of prising reference from

Let’s say that I attended the Obama press conference on April 27, 2006 at the National Press Club in Washington, D.C. The next day I decide to draw what I saw there at a particular moment during the day, from the vantage point of my front row seat.

I produce the first image at right. A week later I set out to draw the same scene, but this time poor memory and lack of common sense conspire against me, resulting in the second image.[13] There is a certain sense in which both pictures depict the very same scene— a particular, real situation that occurred at a certain time and location. In this sense the two images have the same *referent*. On the other hand, there is a clear sense in which they do *not* depict the same thing. The first picture depicts Obama with short hair; the second depicts him with longer, spiky hair. In this sense, the two images have different *content*.

Philippe de Champaigne, “Vanitas” c. 1671

In addition, not all content associated with pictures is genuinely pictorial. In the 17th century painting by Philippe de Champaigne at right, the rendering of the tulips represents life, the painting of the hour-glass represents the passing of time, and that of the skull, inevitable death.[14] Yet none of these symbolic elements figure in the image’s pictorial content, in the sense intended here. Instead, the painting depicts a flower, a skull, and an hour-glass, in a certain arrangement, at a certain time, and under certain lighting conditions, but nothing more.[15]

Finally, the content of a picture is distinct from what it happens to *communicate* to a particular viewer on a particular occasion; rather the picture’s content is what it *depicts* independent of what is communicated.[16] For example, if you write to me asking how my vacation in Nova Scotia is going, I might respond with nothing but the drawing below, thereby communicating to you is that I am passing the summer in happy and idle recreation. But this is not what is depicted: what is depicted is me sailing a boat. However misleading, it is also compatible with the picture on its own that I have had a stressful, miserable experience (shortly after the moment depicted, the boat capsized and my dissertation sank to the bottom of the ocean).

content, including cases of misrepresentation, are discussed by Knight (1930, 75-6), Kaplan (1968, 198-9), and Lopes (1996, 94-8). Cummins (1996, 5-22) observes a parallel division with respect to mental representation. In general, the distinction between pictorial reference and pictorial content is closely aligned with Kripke’s (1977) distinction between speaker reference and semantic reference. A more thorough analysis might reveal that these are species of the same phenomenon. See Burge (2010, 30-46) for an extended discussion along these lines.

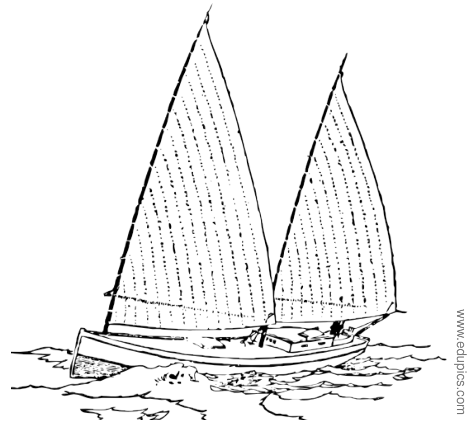
[13]Both drawings are based on the photograph by Mannie Garcia, AP.

[14]See Lubbock (2006).

[15]The distinction between pictorial representation and other kinds of representation by pictures has been made by most authors on the subject; see especially Novitz (1975, §2) and Peacocke (1987, 383). Goodman (1968, 5), for one, appears to reject the distinction altogether.

[16]The distinction between *what is communicated* and *what is depicted* by a picture is borrowed directly from Grice’s (1975, 43-4) distinction between *what is communicated* and *what is said* by a sentence.

In general, what a picture communicates on a particular occasion is independent of what it depicts. What is communicated can exceed the content of picture— as in the case above; or it can fall short— as when a picture is only partially viewed; or the two may be disjoint— when a picture is improperly viewed or interpreted. Of course, if such behavior accompanies *every* act of interpretation, then the content associated with pictures may change or dissolve. The associations of pictures with content inevitably depends on the interpretive practices of the population of artists and viewers in which the picture has currency; I discuss this matter presently. But content is never held hostage to the responses of specific individuals on specific occasions. It is this individual-independent content which I propose the semantic rules associate with pictorial signs.



I have pinpointed the concept of representational content presupposed by the Pictorial Semantics Hypothesis. But should we believe that pictures have such content? The alternative is the null hypothesis that pictures are ordinary, non-representational objects. According to this eliminativist view, pictures carry physical information in the same way that any physical object carries physical information about its causal sources. Thus waves carry information about the direction of the wind and smoke carries information about the location of fire. The eliminativst alleges that this the only sense in which pictures indicate the state of the world; by contrast, I maintain that they carry genuine representational content.

Unlike ordinary objects, representations— objects with representational content— construe the world as being a certain way. As such, they can be evaluated according to whether or not they are correct in their construal. For example, the fact that sentences have content is reflected by the fact that we can evaluate sentence for truth and falsity; truth and falsity are measures of the correctness of a sentence's representational content. By contrast, unadulterated sticks and stones cannot be evaluated for truth or falsity, though they may carry floral and geological information, because they are not bearers of representational content. It is in exactly this sense that pictures are like sentences, and unlike sticks and stones.

While it is admittedly awkward to speak of a picture being "true" or "false", the counterpart of truth-like correspondence with reality for pictures is *accuracy*.[17] It is perfectly natural to evaluate pictures for their relative accuracy and inaccuracy. This reflects the fact, I suggest, that pictures have content. It is in virtue of carrying content that pictures represent the world accurately or inaccurately. The concept of accuracy invoked here plays a central regulatory role in information exchange with pictures. For example, accuracy is the standard of pictorial fidelity which governs high-stakes communicative acts with pictures, like engineering and medical drawing. In such cases, where the information contained in the drawing may form the basis for decisions with life and death con-

[17]Truth and accuracy appear to come apart in two ways. First, pictures and descriptions are accurate and inaccurate, while only sentences are true and false. Second, accuracy comes in degrees while truth is binary. Fodor (1975, 180-1) argues that, even in principle, pictures cannot have truth-values.

sequences, accuracy is the standard which artists characteristically strive for, and viewers expect them to achieve. When the stakes are lowered, it is permissible for artists to infringe upon this standard proportionately. Accurate depiction is simply representation without error. Accuracy in this sense is not the exclusive denotation of the word "accuracy" in colloquial English, but it is one of them.[18]

It may help to clarify what accuracy is not. By accuracy, I do not mean mere *realism*. Realism is a style of painting that approaches illusion: under the right conditions, an observer cannot discern a realist rendering of an object from the object itself. Yet the selective line drawings used by architects, engineers, and medical artists are often perfectly accurate, though hardly illusory.[19] Nor does accuracy imply *precision*. The first drawing of Obama is perfectly accurate in the system of black and white line drawing despite the fact that it is wholly indeterminate with respect to color. It is also accurate even though the lines that compose it are wobbly; this does not mean that the shape of Obama's face is correspondingly wobbly, only that the standards of accuracy determined by this system of depiction are insensitive to a certain level of detail.[20] Nor does accuracy imply *closeness to reality*, in any straightforward sense; a full-scale animated model of Obama is arguably more similar to the man than a black and white line drawing, but both may be perfectly accurate representations. Finally, accuracy does not entail *actuality*. A picture may accurately depict a merely possible scene just as well as an actual one, a point vividly demonstrated by Frank Lloyd Wright's architectural drawing of a *proposed* building plan, with which we began.

For any given state of the world, a picture's content determines whether the picture represents that state accurately or inaccurately. In other words, it determines the conditions under which a picture is accurate or inaccurate. Such *accuracy-conditions* can in turn be used to characterize the content of a picture, just as *truth-conditions* are used to characterize the content of sentences. In §1.4, I'll provide a formal exposition of this same idea; and I'll use this analysis to define pictorial semantics, formally, as a mapping from pictures to accuracy-conditions.

A final concern allows that pictures have representational content, but objects that it is not the sort of content required by a semantics. Instead, semantics essentially involves the definition of *truth*, and counterpart notions such as accuracy fall short of this criteria.[21] While accepting that pictures are neither true nor false, I reject this rigid truth-theoretic conception of semantics. Contemporary semantic theories are keyed to a variety of alternative standards of *correctness*.[22] I simply add accuracy to this roster. I conclude that pictures have representational content, and that this content is a suitable basis for a semantics of pictures.

[18] More or less the same concept of accuracy is widely invoked in discussions of pictorial realism (see, e.g. Lopes 1995; Abell 2007; and Hyman 2006, 194-7). Yet such accounts typically aim to analyze the richer concept of *realism*, taking some form of (non-factive) accuracy as a merely necessary but insufficient condition. Accuracy also corresponds to what Walton (1990, §3.2) calls the relation of "matching".

[19] At any rate, this is *one* sense of the term "realism" among many. See Jakobson (1987 [1921]), Lopes (1995), and Hyman (2006, Ch. 9) for discussion.

[20] See Block (1983, 655-8).

[21] The assumption, widespread among semanticists, is typified by Lewis' (1970, 18) statement: "Semantics with no treatment of truth conditions is not semantics." That said, it is clear that Lewis, if not others, had a quite ecumenical notion of truth. For example, in Lewis (1969), he defines the satisfaction conditions for imperatives in terms of truth. Clearly it is not natural to speak of imperatives as being "true" or "false"; but equally clearly, their content should be characterized in terms of a constraint on how the world must be.

[22] See Hamblin (1958), Kaplan (1999), and Starr (2010) for discussion of a variety of non-truth-conditional approaches.

1.2 Systematicity

I have argued that pictures are bearers of representational content. The Pictorial Semantics Hypothesis further holds that pictures are assigned their content by systematic rules.[23] The thesis that there are systematic mappings from pictures to content is widely doubted. The opposing view is that pictures themselves simply do not make systematic contributions to the information they happen to convey. Like winks or kicks under the table, pictures are thought to impose at most loose and defeasible constraints on interpretation.[24] The rest is left to rational inference and quirks of psychological response; there is no *code* for interpreting pictures.

Tellingly, for the greater part of the twentieth-century, the same kind of anti-semantic attitude dominated philosophical views of language. Ordinary language was thought to be essentially unamenable to systematic analysis. Yet such skepticism ultimately gave way. Following the pioneering work of Chomsky (1965), Montague (1970a, 1970b) and others, theorists of language increasingly warmed to the idea that natural languages are basically rule-governed phenomena. They have even largely converged on the particular proposal that linguistic communication relies on *semantics*— systematic and conventional codes which arbitrarily assign basic meanings to simple linguistic signs, and then derive the complex semantic content of strings of signs, by recursive rules of combination and context sensitivity.[25] This is an explanatory posit, not an article of faith. Humans are able to interpret sentences they have never encountered before, and they are able to do so with extraordinary speed and reliability. Such feats would not be possible, it is thought, if linguistic interpretation were governed solely by open-ended guess work. Instead, these phenomena are best explained by the hypothesis that humans have tacit knowledge of interpretive rules, and they interpret novel sentences by automatically applying these rules.[26] The plausibility of this hypothesis has increased as semantic theories for language have grown ever more detailed and precise.

I propose that essentially the same argument can be made about the interpretation of pictures.[27] The first key observation here is that humans are consistently able to extract content from pictures they have never seen before. For example, a cognitively normal consumer of contemporary global media would have no trouble interpreting the architectural drawing presented at the beginning of this paper, even if she had never seen it before. She would not necessarily be able to identify the particular building depicted, but she would easily recognize the basic three-dimensional spatial and chromatic configuration the picture represents. The fact that we can all interpret unfamiliar

[23]Here and throughout I use the term "systematicity" to denote the property of being systematic, as applied to interpretive schemata. This is distinct from the use of the same term by linguists in discussions of compositionality to denote a specific feature of language in which expressions of the same type can be inter-substituted while preserving interpretability. (Szabo 2010)

[24]Describing the parallel anti-semantic view of language, Stanley (2000, 396) uses the examples of the meaning conveyed by kicking someone under the table or tapping them on the shoulder; Borg (2006, 261) uses the example of raising one's eyebrow.

[25]See King and Stanley (2005) for a canonical statement of this view.

[26]The phenomena described here is commonly termed "productivity." Linguists also argue for the same conclusion on the basis of "systematicity". (See footnote 23.) In either case, the details of the explanation are by no means straightforward. See Lewis (1975), Soames (1989), and Dummett (1996) for three very different perspectives on the matter.

[27]Here I will focus on considerations of productivity for pictorial interpretation. I do believe, however, that it is possible to mount arguments for pictorial semantics from phenomena that correspond roughly to the linguistic concept of systematicity. (See footnote 23.) For example, if a picture is interpretable, so is its mirror image; and certain parts of images can be shifted, rotated, and replaced while maintaining interpretability. I believe the conclusions suggested by this evidence are much the same as those argued for in the main text.

images in this way demonstrates that picture interpretation is not based on a finite list of memorized correspondences. Instead, it stems from some general capacity which may be applied to new particular cases. But what is the nature of this general capacity?

The second key observation is that the interpretation of unfamiliar images is *fast* and *reliable*, despite the fact that both the stimuli and the recovered content are significantly complex. Presented with our initial architectural drawing, normal media consumers would attribute to the image approximately the same spatial interpretation, and each would be able to do so nearly instantaneously. These facts are not easily explained by supposing that interpretation is based merely on open-ended speculation. That is the kind of reasoning we engage when choosing a chess move, or guessing the intentions behind an ambiguous nod. It is characteristically slow and variable, particularly when applied to complex stimuli— nothing like the automatic and consistent responses exhibited in picture interpretation. A better explanation is simply that the mechanisms at work here are governed by general interpretive rules. The speed and reliability with which humans associate unfamiliar complex images with content is explained by the systematic application of these rules. Viewers need not explicitly represent such rules themselves; it is sufficient that they apply certain systematic capacities in regular ways.

It is undeniable that pictures are used in communication; any theory of communication owes an explanation of how this transpires. The considerations reviewed here indicate that at least part of the explanation will be systematic. Positing systematic interpretive mechanisms is the best available explanation for our observations of human interaction with pictures. Of course, this hypothesis is open to empirical falsification. But since this argument is essentially the same as the argument offered for the existence of a semantics of language, the conclusion is just as robust. We should expect to find a semantics of pictures.

The extent and limits of this claim must be understood. First, the claim of systematicity should not be construed too broadly. The preceding argument shows only that a significant portion of pictorial communication is governed by semantic regularities. I do not deny that pragmatic principles play an important role as well. For example, I discussed earlier how an image of a sailboat, sent by mail, would reasonably elicit from its intended audience the judgement that the passengers depicted were especially happy and carefree. But this judgement cannot be warranted by the picture itself, since the passengers themselves are scarcely visible in the image; instead, it is partially justified by the audience's knowledge of the artist's likely intentions, informed by further knowledge of the social norms governing the communicative exchange. That is, the judgement is the result of unconstrained pragmatic reasoning; such imaginative engagement with pictures is ubiquitous. But note that, as is often the case, the pragmatic reasoning on display here depends crucially on first securing that content which is made available semantically. One cannot deduce that the passengers depicted are sailing happily, unless one first appreciates that the image depicts a sailboat, and even more basically, a certain boat-shaped arrangement of edges in space. The extraction of this last dimension of content is strictly determined by the structure of the image itself and the rules of interpretation, and is not the result of pragmatic speculation. It is a substantive and difficult question exactly how much content typically associated with pictures should be counted as purely semantic. My contention here is only that some significant portion is, and in the latter half of the paper I will make this idea precise.

The claim of systematicity must not be understood too rigidly either. An instrumental part of modern natural language semantics was brought on by the development of the theory of CONTEXT-SENSITIVITY. Context-sensitivity arises when the content of a linguistic expression depends on the context in which it is uttered, but in ways that are strictly constrained by the underlying grammar.[28] Perhaps the simplest example in English is the indexical "I", the referent of which varies with speaker, and thus context, but in a manner that is obviously rule-governed. (The grammar requires that "I" refer to its utterer, whoever that may be.) Cases like this illustrate an important moral. The mere fact that the content of a linguistic expression depends in part on context does not imply that such content is not determined by a semantics. Instead, the semantics may associate expressions with meanings which systematically regulate the contribution of context. The same lesson applies to pictorial representation. Given the considerations provided above, we should expect the mapping of pictorial signs to content to be systematic. But as we shall presently see, such content depends in part on context. This does not imply that pictorial representation has no semantics; it suggests instead that the semantics is context-sensitive.

Consider a series of drawings, each produced under different conditions, yet all of which coincidently turn out to be qualitatively identical to the accurate depiction of Obama presented above. One depicts Obama; one depicts a wax sculpture; a third depicts a chance configuration of lines in the sand. In so far as the concepts *Obama*, *wax-sculpture*, and *lines in the sand* are part of— or play an essential role in specifying— the content of each image, such content must be context-sensitive. Since the structure of the image is the same in each case, only the variation in their context of expression can account for the variation in their content. Yet, as the context-sensitivity of language demonstrates, these facts do not imply that the content of each image is *entirely* context-dependent, hence unsystematic.[29] Indeed, it is quite obvious that each scene depicted shares some basic spatial configuration, specified by the structure of the image itself. It is this core component which, I propose, the semantics of pictures directly associates with image structure; the rest is contributed by context.

Following precedents in linguistic theory and recent work in the philosophy of depiction, I propose that a pictorial semantics associates each picture with two levels of content: a primary, EXPLICIT level of content is context-invariant and sparsely geometrical; a secondary, EXTENDED level of content is context-dependent and richly referential.[30] The explicit content of an image is that minimal spatial and chromatic profile of visible surfaces which can be derived from the structure of the image alone. (In the second half of this paper I develop a precise theory of explicit pictorial content.) The extended content of an image elaborates the explicit content with reference

[28]The classical source of this view is Kaplan (1989). See Stanley and Gendler-Szabo (2000) for more recent elaboration.

[29]Balog (2009, §2) presents such an argument against pictorial semantics.

[30]This idea has two sources. The suggestion that pictures have two levels of content— a thin, geometrical level, and a rich, referential level— is due to Haugeland (1991, 73-5), who distinguishes between "bare-bones" and "fleshed-out" pictorial content; this idea is developed extensively by Kulvicki (2006, 59). (Siegel (2006, 2) discusses analogues for perceptual content.) The suggestion that a sign's context-sensitive content can be derived from a thin level of context-invariant meaning, together with context is due to Kaplan (1989). Kaplan theorizes that the grammar associates each lexical entry with a "character" (also called a "standing meaning"), which, together with context, determines "content". Explicit content in my sense is the pictorial analogue of Kaplan's character; extended content in my sense is the analogue of Kaplan's content. Thanks to Matthew Stone for suggesting the term "explicit content".

to individuals, properties, and richer spatial features. It is extended content that is most naturally grasped when we talk about, say, the picture depicting *Obama* as having a certain kind of *hair cut*.

But extended content always depends constitutively on explicit content. For example, it is part of the extended content of the Obama picture that the man's hair has a very specific outline shape, one which we struggle to articulate in words. It is this kind of purely structural feature which is supplied by the explicit content of the image. Explicit content also functions to constrain the way that context may enrich a picture's representational properties. A single image type might accurately depict Obama, or lines in the sand. But the same image could not accurately depict a cube or a plane, no matter the circumstances of its creation. The image's explicit content places hard constraints what the image may be used to accurately depict. It constitutes the core of pictorial representation.

In this paper, I directly defend the systematicity of pictorial semantics only with respect to explicit pictorial content. That is, I defend the claim that there is a context-invariant component to pictorial content which is associated, by systematic rules, with image structure. Any theory of extended content, which takes into account the influence of context, necessary depends on a prior theory of explicit content. This paper is restricted to the more basic task.

Despite the evidence that pictures are systematically mapped to their content, and the appropriate qualifications of this claim, the very possibility of such a mapping is widely doubted. Language, it is thought, is the paradigmatic site of systematic semantics. It is clear that language provides a poor model for pictorial representation, but this fact is taken to impugn the viability of a pictorial semantics. On this view, there can be no semantics outside of language. Three kinds of concern motivate this outlook.

(i) First, it does not seem that pictures have syntax in the manner of linguistic expressions. Semantic rules are typically understood as systematic mappings from syntactic structures to contents. But without syntactic structures to take as inputs, it seems that a semantics of pictures could never gain firm footing. Yet this objection misunderstands the general character of semantics. In the most general sense, it is *structure*, not logical syntax, that is the point of departure for semantics. Thus I grant that pictures do not have a *syntactic* structure like that of linguistic expressions, and they may not be governed by substantive rules of syntactic well-formedness.[31] Yet in the second half of the paper I will develop a view according to which pictures have *geometrical* structure instead. The semantics of pictorial representation associates geometrical structures with representational content using systematic techniques of geometrical projection.

(ii) A second concern has to do with compositionality. Compositionality is often understood as the recursive derivation of the meaning of a complex expression from the meanings of each of a finite number of atomic parts, along with the way these parts are put together. But pictures do not seem to be composed from a finite set of atomic parts. Hence their content cannot be derived compositionally; and without compositionality, it is wagered, there is no semantics.[32] But

[31] For example, it might be that a picture belongs to the system of black and white line drawing just in case it consists of (a) black lines and (b) white regions. This amounts to a rule of well-formedness, but not a substantive one.

[32] Thanks to Jeff King and Mark Baker for pressing this concern.

the worry is unfounded. Compositionality, in its most general construal, is the derivation of the content of a complex expression from its internal structure— it need not require that this structure is decomposable into finitely many atomic parts. In the formalization presented in the Appendix, I demonstrate that the proposed geometrical semantics for picture is fully compositional in this general sense.

(iii) The last objection is purely inductive. The only semantic theory which is well-understood is that of language. There have been no other correspondingly plausible candidates for a systematic account of depiction. And it is fundamentally unclear what such a system of pictorial representation would be. If there were a semantics of pictures, it is thought, such an account would have been established by now. The preceding discussion has amounted to a defense of the very idea of a systematic pictorial semantics. But I have not proven that any such semantics exists. Until a systematic theory of depiction is actually articulated, this last form of skepticism will reasonably persist. It is this concern which the second half of the paper is intended to assuage. The remainder of the present part continues to establish the necessary groundwork for subsequent development.

1.3 Conventionality

I have argued that pictures are assigned representational content by systematic rules. The Pictorial Semantics Hypothesis further holds that these rules are conventional. There are a diversity of interpretive systems for pictures, varying across history and culture, which are selected for use by social convention, just as languages are. The opposing view is that the ability to interpret pictures is simply a manifestation of the biologically fixed abilities which make up human visual perception. If there were a theory of pictures, it would be a byproduct of the science of vision, not a topic of independent interest.[33] On this view, conventional knowledge plays no role in the interpretation of pictures.

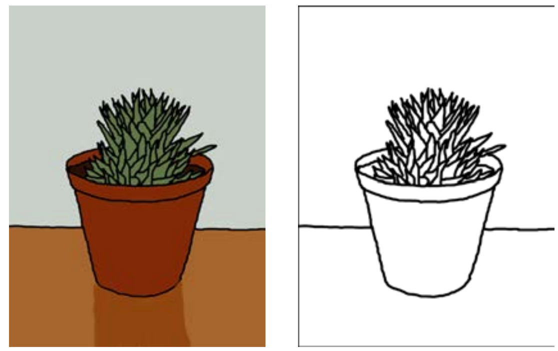
Despite its naïve plausibility, this deflationary hypothesis is now widely discredited. As early as 1921, the linguist Roman Jakobson observed, "it is necessary to learn the conventional language of painting in order to "see" a picture, just as it is impossible to understanding what is said without knowing the language." (1921, p. 21) Since then, theorists of pictorial representation have documented a diversity of pictorial conventions, or SYSTEMS OF DEPICTION.[34] Such systems crucially mediate between pictures and the informational content they carry.

Consider, by way of illustration, the pair of images below. The first contains spatial as well as chromatic information about the scene featuring my plant— e.g. that the *cone-shaped* object is a particular shade of *brown*. Part of what determines its chromatic content is of course the color of the pigments in the picture itself. If these colors had been different, the picture would have had a different content. By contrast, the second image is a black and white line drawing of the same scene. It has the same spatial content as the first drawing, but contains no chromatic information about

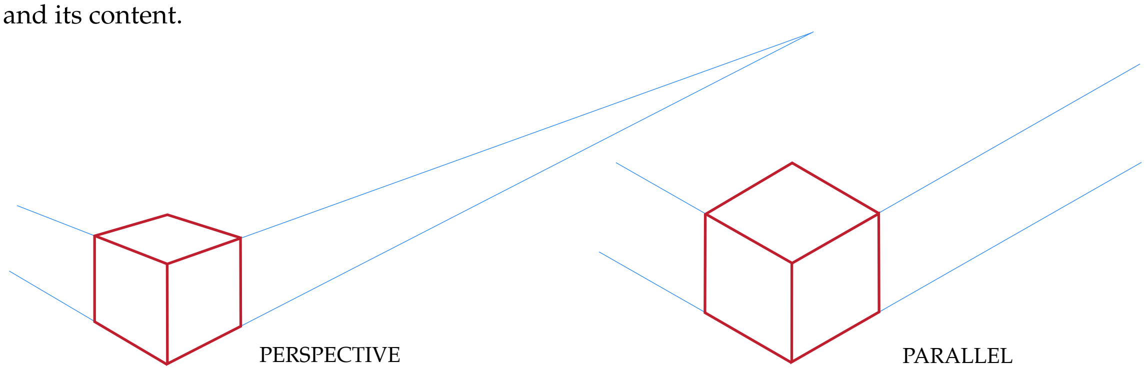
[33] This need not imply that seeing a picture of a scene is the same as seeing the scene itself. After all, black and white drawings look nothing like the scenes they depict. Instead, it is held that the mechanisms which ground picture interpretation are fully determined by (but not necessarily the same as) those which ground normal visual perception.

[34] Three excellent books on the topic are *Art and Illusion* (Gombrich 1960), *Varieties of Realism* (Hagen 1986), and *Art and Representation* (Willats 1997). The term "system of depiction", or a close cousin, is widely used (see, e.g. Manns 1971, 286; Newall 2003, 384; Abell 2009, 222).

the scene it depicts. If it *were* a color drawing, it would depict a plant, pot, and desk which were all paper-white; plainly this is not the case. So while both pictures have content, the relationship between the surface properties of the picture and its content is different in the two cases. The explanation for this is that different systems of depiction, like line drawing versus color drawing, determine different ways of mapping pictures to content.



Differences between systems of depiction are not limited to variation in the treatment of color; they differ in their handling of line, shading, and texture as well. Importantly, they also determine divergent treatments of pictorial *geometry*. Consider the two drawings below; each depicts a cube, but in a different way. In the drawing on the left, *converging* lines in the picture are used to depict the parallel edges of the cube. In the drawing on the right, *parallel* lines are used to depict the parallel edges of the cube. The first is rendered in a PERSPECTIVE system, the latter in a PARALLEL system.[35] If the drawing at right *were* in perspective it would depict an irregularly shaped solid— not a cube at all. Thus, like line and color drawing, these two systems determine different relations between features of the picture surface and its content.



Far from artificial constructs, these systems of depiction and a vast diversity of others have been widely used across history and culture. Classical Chinese painting, for example, employed a type of parallel depiction; another dominated ancient Egyptian imagery, with very little variation for over 2,000 years; meanwhile, Western post-Renaissance depiction has typically been presented in perspective.[36] Notably, parallel depiction systems like those used by classical cultures are the norm in contemporary engineering and architectural drawing. All such systems exhibit a characteristic stability, whereby artists and viewers may continue to exploit the same system across time and in varying contexts.

These facts cannot be explained by appeal to the normal workings of the perceptual system alone. The activity of the perceptual system is widely recognized to be cognitively "impenetrable".

[35] The class of parallel systems includes the more familiar and specific systems of oblique, isometric, and orthogonal projection. Orthogonal projections include plan and elevation depictions common in architecture and engineering. The drawing above is an isometric projection.

[36] On classical Chinese depiction, see Willats (1997, 55). On classical Egyptian depiction, see Hagen (1986, 168-72).

(Pylyshyn 1999) That is, its internal processes are in an important sense automatic; they cannot be modified through the interference of rational inference or conscious decision making. The visual system is also believed to have essentially the same structure in all human populations. By contrast, the use of systems of depiction varies widely across human populations, and depends heavily on contingencies of social context. Thus the interpretation of pictures cannot be explained by appeal to biological vision alone.

This conclusion is compatible with the apparent fact that those systems of depiction used by human societies depend in crucial ways on the workings of the human visual system. The interpretive computations required by systems of depiction are broadly of a kind with those performed by the human visual system in the course of normal vision. This idea is made concrete in the semantic theory developed later. Because of their perceptual origins, the space of humanly realizable systems of depiction is indeed constrained by universal features of human vision. Not anything can become a human system of depiction.

Yet the conventionality of pictorial representation frees it from the strictures and idiosyncracies of vision. Systems of depiction introduce expressive devices never found in perceptual experience. These include outlines, alternative color schemes (like black and white, or negative color), color coding, and alternative projection systems (like parallel projections or map projections), to name just a few. Such departures from normal vision demand interpretive familiarity. Systems of depiction also determine objective standards of accuracy that do not depend on the visual impairments or special abilities of individuals. This facilitates reliable communication. In this way, systems of depiction are both conventional crystalizations of visual perception and extensions of it.[37]

Thus systems of depiction are the pictorial analogues of languages: culturally specific interpretive rules for encoding content in signs. Are systems of depiction therefore conventional in the same way that languages are? They are. Although systems of depiction have a very different internal architecture from languages, the use of systems of depiction in human communication is conventional. More carefully: there are infinitely many possible systems of depiction, considered as abstract interpretive rules; among these, a few are selected by social convention for use in human communication, just as languages are.

Following Lewis' (1969) influential account, conventions of interpretation, such as languages, arise when a population wishes to communicate, but more than one methods of communication is available for use. In order to successfully communicate, the population must collectively coordinate on one such method for repeated use, thereby establishing a convention. Systems of depiction are conventional in exactly this way. Consider a concrete example: the engineers in a construction firm must share drawings amongst themselves while designing and building a bridge. It is important to everyone that all the drawings produced within the firm be accurate and be accurately interpreted, since inaccuracy in production or interpretation could have disastrous effects. The only reliable and efficient way to achieve this goal is to establish a standard system by which drawings are produced and read. That is, the engineers must collectively establish a convention of producing and interpreting drawings according to a given system of depiction.

[37] See Gombrich (1960, 3-30), Kennedy (1974, 24-7), Lopes (2009, 18-9), Durand (2002, §4), and DeCarlo and Stone (2010, 5).

The same phenomena is repeated, at a larger scale, in contemporary media. For example, the use of perspective line drawing is a global convention. Here, the coordination among artists and viewers is tacit and disparate. Nevertheless, the basic considerations are the same: all seek to efficiently communicate approximately accurate information; yet there are always multiple ways this might be achieved; communicative success is ultimately realized by coordinating on a stable interpretive convention for pictures. Of course, contemporary media relies on a great variety of systems of depiction. Just as many human societies are multi-lingual, human societies can and do use a multitude of systems of depiction. Those which are used recurrently become conventions.

In spite of the considerations reviewed here, many have thought that systems of depiction *could not* be conventional, even in principle. Paralleling the debate about systematicity, objections here tend to take languages as the paradigm case of conventionality, and argue that, because of their disanalogies with language, systems of depiction could not be conventional in the same way. Dispelling such concerns will help elucidate the central claim of this section. I focus on three objections in particular.

(i) If systems of depiction were conventional, it is thought, they would be arbitrary in the way that languages are. Yet while words bear a merely arbitrary relationship to their content, pictures bear a much more intimate and direct relationship to theirs. Thus systems of depiction cannot be conventions. This objection trades on a simple but widespread confusion between the sort of arbitrariness which is a feature of all conventional behavior, and the sort of arbitrariness associated specifically with language.

It will be helpful here to recall a key feature of Lewis' analysis of convention. Lewis observed that for a regularity of behavior within a population to count as a convention, there must be some alternative regularity with which the population *could* have conformed, while still realizing their dominant needs. Part of what makes a regularity a convention, then, is that it reflects a *choice* (not necessarily a conscious one) on the part of the group, between mutually acceptable options. For example: driving on the left is a convention in some nations, in part because drivers *could* have collectively driven on the right instead, while still realizing their dominant need to pass each other while driving.

The sense in which *all* conventions are arbitrary is simply this: a population which conforms with a given convention could have conformed with a different convention, while still realizing their dominant goals. The use of, say, perspective line drawing by some population is a convention, in part because that population could have used parallel projection line drawing for the same purposes instead; in that sense the use of perspective line drawing by that population is arbitrary. In the same way, the use of English by some population is a convention, in part because the population could have used Shona for the same purposes; in that sense the use of English is arbitrary. But this kind of arbitrariness is not unique to systems of symbolic representation.[38]

The other kind of arbitrariness *is* unique to systems of symbolic representation. This is the kind displayed by the arbitrary link between the word "tree" in English and the property it denotes of being a tree; there is no intrinsic connection between the sign and its content. The same kind of arbitrariness is *not* displayed by the link between a picture of a tree and the tree it depicts.

[38]Eco (1979, 189-92) is one of the few authors to make this point explicitly.

Since there are non-symbolic systems of representation, such arbitrariness is not a necessary feature of conventional systems of representation. And it is not even sufficient. As Fodor (1975, 178) has observed, if there are mental symbols, they bear an arbitrary relation to their content, but the relation is not determined by convention.[39]

Thus it is essential to distinguish one kind of arbitrariness, *in a group's choice to use a particular system of representation*, from arbitrariness *in the relation between a sign and its content within a system*. The former is a necessary condition on conventionality; the latter is specifically a feature of language, compatible with, but not necessary for conventionality. "Arbitrary" here has simply acquired two related but independent meanings. Public systems of depiction and public languages are both selected for use, by populations, conventionally. But of the two, only public languages rely on arbitrary pairings of signs and contents. A simplified example may help to make this point more vivid.

Suppose you and I must communicate our credit card numbers in a crowded room. To guard our privacy, we must come up with some kind of code. Here are four codes we might use:[40]

+1 Code	*+2 Code*	*A-List Code*	*B-List Code*
"1" means 0	"2" means 0	"9" means 0	"2" means 0
"2" means 1	"3" means 1	"3" means 1	"5" means 1
"3" means 2	"4" means 2	"4" means 2	"4" means 2
and so on...	and so on...	"7" means 3	"8" means 3
		"5" means 4	"9" means 4
		"2" means 5	"6" means 5
		"8" means 6	"1" means 6
		"1" means 7	"3" means 7
		"0" means 8	"7" means 8
		"6" means 9	"0" means 9

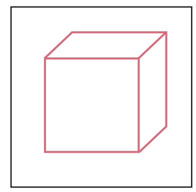
The +1 and +2 codes exploit simple computations to disguise their message; other such codes could employ other once-place mathematical functions like multiplication or exponentiation by a constant. The A-List and B-List codes employ distinct, arbitrary pairings between numerals. There are as many such possible codes as their are possible distinct sets of pairings. Let us suppose that all four codes described here are sufficiently secretive for our purposes.

If the problem of communicating with codes is a recurrent one, it is convenient to establish a convention; the convention dictates which of the infinitely many possible codes (including the four above) we are to select for use on subsequent occasions. In the first sense, the selection would be arbitrary. Because all four codes suit our purposes equally well, the choice to use the +1 code instead of the +2 code or the A-List code is an arbitrary choice. But now consider that the A-List and B-List codes are *arbitrary* in a way that the +1 and +2 codes are not. They rely on arbitrary pairings of input and output, whereas the others rely on mathematical laws. Clearly, then, arbitrariness in the selection of a code for use is quite a different thing from arbitrariness in the structure of the code itself.

[39] I discuss this kind of arbitrariness at greater length in Greenberg (2011).

[40] Here I put the message expressed in quotation marks, to distinguish it from the numeral thereby indicated. Strictly speaking, it is probably most apt to think of both relata of the code as numerals, not numbers.

The lexicons of natural language are the analogues of the A-List and B-List codes. They are conventions of interpretation which are based on arbitrary pairings of signs and contents. Systems of depiction are the analogues of the +1 and +2 codes. They are conventions of interpretation which are based on law-like associations between signs and contents. Although I have said little thus far about the internal structure of systems of depiction, in the second half of this paper I develop a theory according to which they are based on rules of geometrical projection. For present purposes, such projective rules are merely complex cousins of the +1 and +2 codes. The general moral is this: conventionality requires that the agents exploiting the conventional rule have made an arbitrary selection of that rule over another. But conventionality makes no claim about the nature of this rule. It may be arbitrary, like a lexicon, or law-like, like a geometrical projection. Systems of depiction, as it happens, are often deployed in communication by convention.



(ii) A second objection holds that, for all we have just said, languages must be *more* conventional than systems of depiction. The evidence is this: pictures produced according to one system of depiction can typically be successfully interpreted by viewers familiar only with other, even very distant systems.[41] For example, a viewer familiar only with systems based on the geometry of *perspective* projection would be able to recover the spatial information from the picture at right, even though it is the product of a system based on *parallel* projection. By contrast, sentences from one language are not generally comprehensible to speakers of another language. Thus languages must be more a product of convention, and systems of depiction less so.

The argument relies on an implicit false premise: although viewers may manage to recover the *crude* content of pictures from unfamiliar systems, it is never possible to determine with certainty the *exact* content of a picture without acquaintance with the relevant system. Thus, without knowing whether the picture above belongs to a perspectival system, or a system with some alternative geometry, it is impossible to tell whether the object represented is in fact a cube, or a related but irregular solid. Conventional knowledge is required, after all, for correct interpretation.

What the objection *does* reveal is that the space of humanly implementable systems of depiction is narrower than the space of humanly implementable languages, restricted as it is by the universal structure of the human visual system. As a consequence, human viewers can generally make educated guesses about the interpretation of pictures in unfamiliar systems. Yet this difference between languages and systems of depiction is only one of degree. Genetically determined neural anatomy put severe constrains on the space of languages that can be used by humans, as well.[42] In both cases, restrictions are imposed by the biological structure of the interpretive anatomy.

Ultimately, it is immaterial to the conventionality of a given system of representation whether it belongs to a wide or narrow space of humanly realizable alternatives. The conventionality of a given system requires only that there exist *some* alternative. Driving on the left, for example, may be a convention, even when there is only *one* possible alternative which satisfies the goals of population— driving on the right. Drawing in perspective, like driving on the right, or speaking French, is a matter of social convention.

---

[41] See Carroll (1999, 42-4).

[42] See Pinker (1995).

(iii) A third and final objection, due to Lopes (1992, §3-4; 1996, §6.6), contests the claim that systems of depiction are conventional, on the grounds that the choice to use one system of depiction over another is rarely arbitrary. For example: drawings from systems of black and white depiction and those from systems of color depiction reveal different features of their subjects, and require different technologies to produce. The choice to use one kind of system or the other in a given communicative exchange is invariably motivated by the practical demands of the task at hand. We might specifically wish to communicate color information, or specifically wish to suppress it. But, it is argued, whenever the choice to use one system or another has a specific rationale that use is no longer based on convention.

The objection rests on the principle that conventions only arise when the population in question is truly indifferent between alternative regularities in behavior, so that the choice between them is entirely arbitrary. But the principle is false. It is true that, for a regularity in behavior to be a convention, the population must select that regularity from a space of alternatives. But the selection need not be unmotivated. By way of illustration, consider a population for whom, through a quirk of genetics, driving on the left side of the road is discomfiting, while driving on the right side of the road is relatively pleasurable. It is a matter of common knowledge throughout the population that this is so. Nevertheless, if all members of the population drove on the right, and expected all others to do the same, it would still be a convention, despite the fact that it is the mutually preferred alternative. It would be a convention, in part, because all would have been prepared to drive on the left, if necessary, in order to pass each other on the road. Thus a regularity in behavior can be a convention, despite the fact that it would be the preferred behavior quite apart from convention; also a regularity can be a convention even if it would be dispreferred. Conventionality does not require indifference.

In his analysis of convention, Lewis (1969) is sensitive to these facts. Rather than requiring that a populations be indifferent between alternative regularities, Lewis imposes a subtler, and more plausible requirement. For a regularity to be a convention, it must be one among a set of alternatives which are such that each member of the population prefers conformity to that regularity *on the condition that everyone else conforms*. Thus, I prefer to speak English if everyone else speaks English; I prefer to speak Shona if everyone else speaks Shona. This condition can be met, even if the alternatives are preferentially ranked. All things being equal, I may prefer to drive on right; but this preference is trumped by the desire to drive in the manner of everyone else. This is sufficient to qualify the regularity for conventionality.

In the original case of depiction, the dominant goal of the typical communicative exchange is to communicate accurate information using a picture. Thus, while it may be that, all things being equal, I prefer to produce and view color drawings, I prefer most to use the same system of depiction that everyone else is using. Under such conditions, a system of black and white depiction or one of color depiction could become the convention. Thus the conjecture that the use of any given systems of depiction is often or always motivated by practical concerns does not impugn its conventionality.

Together, the last three sections constitute my defense of the Pictorial Semantics Hypothesis. I have argued: (i) that pictures are bearers of representational content; (ii) that the mapping between pictures and their content is systematic; and (iii) that there are a diversity conventional systems according to which pictures are associated with their content. The Pictorial Semantics Hypothesis can now be more precisely construed as the hypothesis that every actually occurring system of depiction determines a systematic and conventional mapping from pictures to their content. The aim of the next section is to render this claim with sufficient formal precision that we may reliably evaluate particular theories of pictorial semantics.

1.4 Accuracy-conditional pictorial semantics

The aim of this section is to give formally explicit expression to the Pictorial Semantics Hypothesis, by defining the concept of pictorial content adumbrated above in terms of its accuracy conditions relative to a system of depiction. This accuracy-conditional framework will set the stage for the assessment and development of specific theories of pictorial semantics in the remainder.

We begin by developing a formal model for depictive content that abstracts away from the distinguishing features of different systems of depiction.[43] There are a number of ways to proceed here, but following Blumson (2009b), a natural course is to extend the popular possible-worlds theory of linguistic content to the case of pictures.[44] On this view, since a picture represents a certain way things might be, its content can be modeled by the set of possible worlds so represented.[45] But this is too crude: a world could be just as a picture represented it at one *time*, but not at another. Similarly, it could be just as a picture represented it at one *location*, but not at another. It is therefore inappropriate to think of pictures as representing the way the world is across the board; instead we should think of them as representing worlds at particular times and locations.[46] (For now I choose to ignore the further matter of a picture's *angle of view*. I will return to this issue later once we have developed a richer understanding of the mechanics of perspective.) We can think of these particular locations and times within worlds as *scenes*. Pictures do not, as adverted before, represent the entire world as being a certain way— they represent a more limited scene as being a certain way. Thus the content of a picture is modeled as a set of scenes. We may define a SCENE simply as a ⟨*world, location, time*⟩ triple, what is sometimes called a "centered world."[47]

Depictive content is modeled by *sets* of scenes, rather than individual scenes, because pictorial

[43]There are two sorts of theory of content. One seeks to place purely *formal* constraints on the structure of content, but remains silent on the metaphysical nature of the objects realizing these constraints. This is the norm in linguistic theory. (See Lewis (1970, 32) for a brief but apt defense of this norm.) The heroic few (e.g. King 2007) take on both the formal and metaphysical questions. In this paper I am concerned only with the formal characterization of pictorial content.

[44]For a canonical statement of the possible-worlds account of content, see Stalnaker (1984). The application of this theory to pictorial content has been proposed by Howell (1974), Ross (1997, 17-21), and Blumson (2009b). Alternatively, we coud define the content of a pictures as a complex property (Lopes 1996, 111) or as a structured situation (Malinas 1991, 276). Further proposals might be usefully adopted from the theory of linguistic content, perceptual content, or the content of mental imagery.

[45]Does the existence of "impossible" pictures like the Penrose triangle or the illustrations of Escher require that pictorial content also include *impossible* worlds? (See Huffman (1971) and Sorensen (2002) for examples.) It is known that all familiar impossible pictures can be divided into possible parts. (Kennedy 1974, 149-50; Sorensen 2002, 346) On the strength of this observation, I deny that impossible pictures have content, while allowing that their possible parts are individually contentful, though mutually incompatible. Thus the pressure to countenance impossible worlds is diffused. Blumson (2009b, §7), for one, opts instead to include impossible worlds in pictorial content.

[46]See Blumson (2009b, §2).

[47]The original idea is due to Lewis (1979). For our purposes, it doesn't matter much how we think about possible worlds, though conceptions of worlds as constructed from linguistic objects may introduce certain difficulties.

representation is always indeterminate.[48] In the system of black and white line drawing, for example, the picture of Obama does not represent Obama's suit as being any particular color. So the picture is compatible with worlds in which the suit is blue, as well as worlds in which it is grey or red. While color systems are not indeterminate in this particular way, nearly all systems are indeterminate with respect to *occluded surfaces*— that is, those surfaces obscured by some more nearby surface. Thus the picture of Obama does not represent Obama's *back* as being any particular way; it is compatible with worlds in which Obama has wings, occluded from view, and worlds in which he doesn't. In such cases the picture fails to specify a complete description of the scene before it, hence the characterization of its content as a set, rather than a single scene.

Though the content of a picture may always be modeled by a set of scenes, *which* set of scenes must depend on the operative system of depiction. For example, the set of scenes compatible with the content of a color drawing is more restricted chromatically than those compatible with its black and white counterpart. A quite natural suggestion is to define pictorial content in terms of DEPICTION— understood as a two-place relation between a picture and a scene, specified relative to a system of depiction.[49] The idea is that a single picture may depict a range of possible scenes— for example a single black and white line drawing may depict a chromatic variety of scenes— but the placement of lines in the picture ensures that every scene the picture depicts has the same basic configuration of objects in space. Taken together, the set of scenes which a picture depicts in a system defines the picture's content in that system.[50]

My only complaint here is that the term "depicts" can be used ambiguously to track either pictorial reference or pictorial content. What we need to specify is a depictive relation that holds between all and only the scenes which conform with a picture's content, independent of its intended referent. This can be isolated by focussing on the *accurate* or *successful* depiction of a scene. Thus the spiky-hair image of Obama *accurately* depicts all and only scenes with spiky-hair-shaped objects of the appropriate sort— though none of these scenes are the actual referent of the picture. The definition of pictorial content can then be expressed unambiguously, for any picture P and system of depiction I:[51]

(2) The content of P in I = the set of scenes S such that P accurately depicts S in I.

The content of a picture, according to this view, is its *accuracy conditions* in a system. Note that although accuracy can come in degrees, my interest here is in *maximal* or *perfect* accuracy. The content of a picture is the set of scenes it perfectly accurately depicts.[52] Here and throughout, by

[48]See Lopes (1996, §6.1) and Ross (1997, 18).

[49]A different approach, pursued by Ross (1997) and Howell (1974) is to characterize pictorial content indirectly, by considering what it must be in order to satisfy true sentences about pictures. This kind of consideration has its place, but my primary interest here is in pictures, not language about pictures. Still another strategy is developed by Malinas (1991, 288-9), who defines picture content in terms of what "queries" a picture answers with "yes", "no", or "neither". It is hard to evaluate this proposal, because such terms are so alien to the concepts we ordinarily use in thinking about pictures.

[50]This way of defining content is implicit in Blumson (2009b).

[51]An accuracy-conditional definition of content is briefly considered in Haugeland (1991, 78) and Blumson (2009b, §3). Appropriately enough, parallel accuracy-conditional accounts of content has been proposed for perceptual experience. (Siegel 2009, §2; Burge 2010, 34-42)

[52]It seems to me that a definition of perfect accuracy is *all* one needs to specify the content of a picture. Part of the content of the spiky-hair image of Obama is that Obama has precisely *that* shape of hair— not close-cropped hair, not even hair which is slightly less spiky, nor hair which is slightly more spiky. For any hair-shape not perfectly accurately

"accurate depiction" I mean perfectly accurate depiction, or very near it.

The concept of accuracy at play here is that which is required to define pictorial content in the way specified by (2). But it is more than that: it is the same concept which figured in my initial defense of the idea of pictorial content in §1.1. As I argued there, not only do we easily make direct and robust judgements of accuracy, it is the standard of fidelity to which artists aim, and viewers expect them to achieve, particularly in high-stakes contexts of communication with pictures. In short, accuracy is the basic measure of success for pictorial representation.

I have defined the content of a picture in a system as the set of scenes or "conditions" it accurately depicts in that system. I thereby join a formidable rank of philosophers and linguists to characterize the content of a variety of representational kinds in terms of some type of *correctness* conditions.[53] Most prominently, natural language semanticists in the twentieth century have overwhelmingly specified the content of assertoric sentences in terms of *truth*-conditions. The same kind of approach has been fruitfully extended to questions, commands, and even perceptual experiences.[54] In general, observing the connection between content and conditions of correctness has been the key to the development of formal semantics as a precise and progressive research program; here I follow this precedent.

The Pictorial Semantics Hypothesis can now be formally articulated as the claim that, for every system of depiction, that system determines a systematic mapping from pictures to accuracy conditions. A theory of pictorial semantics for a given system is therefore a theory of accuracy in that system. Given a system of depiction I, such a theory must complete a schema of the following form, for any picture P in the system I, and scene S:[55]

(3) P accurately depicts S in I if and only if __________________.

Ultimately, the empirical plausibility of any semantic theory depends on the specific predictions it makes in particular cases. The primary import of the accuracy-conditional schema presented here is to provide a standard format by which the predictions of such theories can be expressed, and their success rigorously measured. It provides the basic framework for the development of a pictorial semantics in what follows.

1.5 The scope of inquiry

In preceding discussion I have defined and defended the hypothesis that there are pictorial semantics, systems of depiction which determine systematic and conventional mappings from pic-

represented by the picture, that shape is not part of the content of the picture. Methodologically, this is good news, for to specify the representational content of a picture, a semantic theory need only supply necessary and sufficient conditions for perfect accuracy without involving itself in the dalliances of gradability.

[53]See Stalnaker (1998), Kaplan (1999), Gunther (2003, 4-6), and Burge (2010, 38-9) for discussion of the connection between content and correctness conditions. Once again, I mean to include in this group those theorists who refuse to *identify* some signal as its correctness conditions— for example defenders of Russellian and Fregean theories of content— but who nevertheless accept that their preferred notion of content must determine correctness conditions.

[54]The truth-conditional approach has dominated twentieth-century semantics. Contemporary theorists in the dynamic tradition have added to this a variety of "update" conditions; see, for example, Groenendijk and Stokhof (1991) and Veltman (1996). In the realm of non-assertoric sentences, Hamblin (1958) proposes that the content of questions be defined in terms of "answer-hood" conditions; see Groenendijk and Stokhof (1997, 24) for further discussion. And see Starr (2010, Ch. 4) for a recent non-truth-conditional analysis of imperatives. Perceptual content, for its part, is often defined in terms of accuracy conditions. (Siegel 2009, §2)

[55]This schema is based on Tarski's (1944, 344) T-schema for truth.

tures to content. In the remainder I take up the question of what sort of semantics pictures actually have. While the final aim of any pictorial semantics is to complete the semantic schema articulated above, I will limit the scope of the following inquiry in two important ways.

First, philosophers of language distinguish between FOUNDATIONAL and DESCRIPTIVE semantic theories, each necessary components of a complete theory of content.[56] Foundational semantic theories attempt to explain, from the metaphysical ground up, how inanimate concrete objects come to have representational properties.[57] Such theories are characteristically concerned with how an intelligent agent's interactions with an object, in a social context, manage to imbue that object with representational content. Descriptive semantic theories, by contrast, simply assume a class of objects with representational properties, and aim instead to explain why these representations have the content that they do. Such theories are the norm in linguistics, aiming only to articulate the formal rules which map linguistic signs to their content. This is typical in the philosophy of depiction as well, where the dominant theories, such as resemblance and perceptual accounts, are primarily distinguished by their approaches to the descriptive analysis of pictorial content.[58] My ambitions in this paper are the same. I'll have nothing to say about what separates accidental marks on a piece paper from genuine pictures. Instead, I'll *begin* with the class of pictures, and proceed directly to a descriptive account of the rules by which they are associated with their content.

Second, as I indicated above in §1.2, the discussion that follows is only concerned with the context-invariant, or explicit, component of pictorial content. While a final theory of pictorial representation must also contend with the context-sensitive, extended content of pictures, any such theory necessarily depends on a prior account of explicit content. So it is there that I begin. Henceforth, when I speak of "accurate depiction" I mean accurate depiction with respect to explicit content alone. In this sense, the same picture may accurately depict Obama, a wax sculpture, or lines in the sand.

Eventually, a theory of pictorial semantics must come to grips with both the foundational and the context-sensitive aspects of depiction. But we are still very far from that achievement, and there is a danger of seeking completeness at the cost accuracy and detail. I propose instead to purse a rigorous account of the central, context-invariant semantic architecture of certain common systems of depiction. I hope this will provide the firm bedrock which will ultimately be necessary for a systematic account of all aspects of pictorial representation.

2 The Geometrical Projection Theory

The general thesis of this paper is that there is a semantics of pictorial representation. The chief obstacle to this claim is the scarcity of theories of pictorial representation which are at once sufficiently systematic and compelling to qualify as promising theories of pictorial semantics. In this section, I'll present a largely novel account of depiction which aims to stand apart from the

[56]See Stalnaker (1997, 535-6) and Stanley and Gendler-Szabo (2000, 220-6).

[57]This is the sort of theory taken up by Putnam (1975; 1981), Field (1978), Dretske (1983), Stalnaker (1984), and Fodor (1990a, 1990b). See Loewer (1997) for a useful overview.

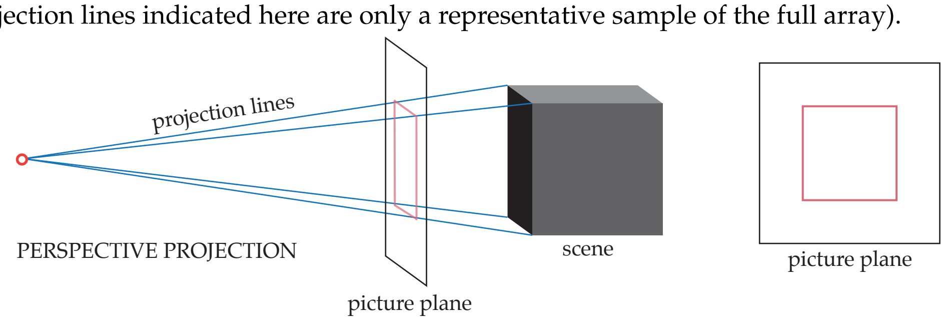
[58]Abell and Bantinaki (2010) provide a clear discussion of this point, in the context of philosophy depiction, including a useful survey of foundational theories of depiction from the literature.

existing proposals by fulfilling this mandate. As defined in §1, a pictorial semantics specifies the conditions under which a picture accurately depicts a scene relative to a given system of depiction. A natural hypothesis is that these conditions involve some complex relation of spatial and chromatic "fit", such that, when the relation holds between a picture and a scene, the picture accurately depicts the scene. Inspired by the analysis of images in projective geometry and computational perceptual science, the theory I'll defend here holds that, in any system, the relation in question is one of *geometrical projection*: a picture accurately depicts a scene just in case the picture can be derived from the scene by a method of projecting three-dimensional scenes onto two-dimensional surfaces.

This proposal is developed in three parts: §2.1 briefly introduces the key concept of geometrical projection; §2.2 then articulates what I call the *Geometrical Projection Theory*; finally §**??** distinguishes the theory from the main alternative accounts of depiction. In §**??** the theory is elaborated in greater detail for the particular system of perspective line drawing.

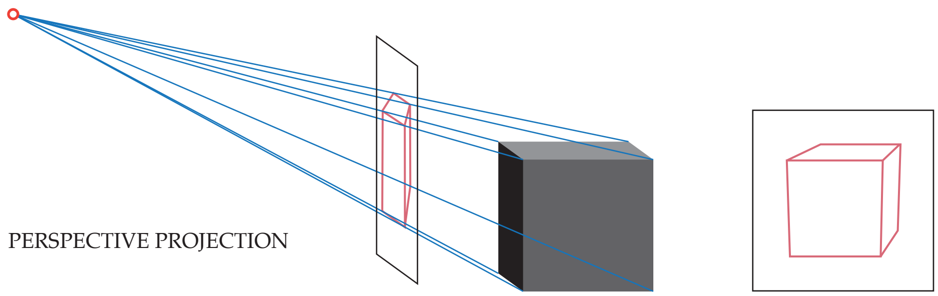
2.1 Geometrical Projection

GEOMETRICAL PROJECTION is a general method for transposing three-dimensional scenes onto two-dimensional picture planes, much the way a flashlight may be used at night to project the shadow of an extended object onto a flat wall. The method works by first defining an array of PROJECTION LINES which connect a scene to a picture plane according to a rule of projection; these lines are then used to map spatially distributed features of the scene to surface features of the picture. A simple example is illustrated below, with the resulting picture plane revealed at right.[59] (Note that the projection lines indicated here are only a representative sample of the full array).

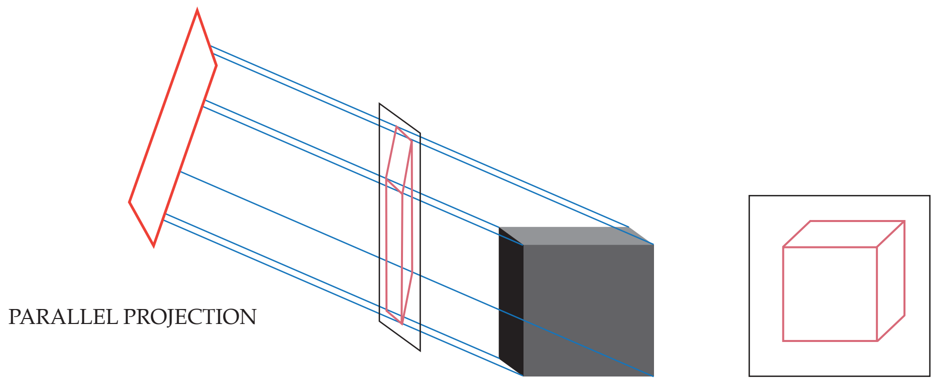


In this particular method of projection, known as PERSPECTIVE PROJECTION, all the projection lines are defined by a single point of convergence (the ∘ above). This point can be shifted, with the effect that new features of the scene are revealed in the projected image. The result is a different *projection* from the one just illustrated, but the *method* of projection— perspective— remains the same.

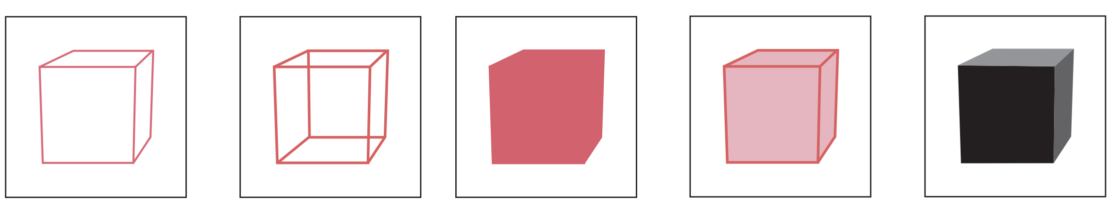
[59]The diagrams of this paper are all drawn according to a highly imprecise system of depiction. The information they convey is impressionistic, not exact.



Perspective projection was first fully codified by artists and scholars of the Italian Renaissance, but it is only one of a great many other methods of projection. The diagram below illustrates a version of PARALLEL PROJECTION, in which the projection lines, rather than converging on a single *point*, are all perpendicular to a single *plane*, hence parallel to one another. The resulting image is subtly but visibly distinct from the perspective projection.[60]



These examples illustrate how methods of projection may vary with respect to the configuration of their projection lines, hence distinct *structural* relationships between the scene and the projected image. But systems can also be distinguished according to *which* particular features of the scene are mapped to the picture plane, independent of the structure of the mapping. For example, in all of the drawings considered above, *visible edges* in the scene were mapped to *lines* on the picture plane. But there are many other possibilities: in "wireframe" projections, *all* edges in the scene are mapped to lines in the picture. In methods of color projection, colors in the scene are mapped to suitably related colors in the picture. These and other variations— by no means exhaustive— are illustrated below, in combination with the distinctive structure of perspective projection.[61]



visible edges ⇒ lines | all edges ⇒ lines | surfaces ⇒ red | visible edges ⇒ lines, visible faces ⇒ pink | colors ⇒ colors

In §3 I will explain the method of projection for perspective line drawings in considerably

[60]The projection illustrated below is an oblique projection, a species of parallel projection.
[61]This explanation of geometrical projection, along with the notation for feature mappings, is drawn from Willats (1997).

more detail, and in the Appendix the definition is explicitly formalized. For the present, I hope the preceding illustrations are sufficient to provide the reader with a working understanding of the general notion of geometrical projection. With this in hand, we are now in a position to examine this paper's specific proposal for a semantics of pictures.

2.2 The Geometrical Projection Theory

What is the relation that characteristically holds between a picture and a scene it accurately depicts? As we'll see, philosophers of art have considered a variety of answers to this question. The orthodox view is that the relation is one of similarity with respect to a carefully constrained set of spatial and chromatic properties. A radical reaction is that the relation is utterly arbitrary, determined by a library of picture-scene associations. A very different approach holds that the relation is defined in terms of the perceptual effects of the picture on a suitable viewer. The theory on offer here diverges markedly from all of these, holding instead that the relationship in question is simply that of geometrical projection.

Officially, the Geometrical Projection Theory has two components: the first describes the relationship between systems of depiction and methods of geometrical projection, while the second defines accurate depiction relative to a system in terms of projection.

(4) ***Geometrical Projection Theory***
  (a) Every system of depiction determines a unique method of projection.
  (b) A picture accurately depicts a scene in a given system if and only if the picture can be derived from the scene by the projective method characteristic of that system.

I have in mind a formally precise rendering of this hypothesis, at least for the particular system of perspective line drawing, which I will develop in the next section. But for now, my aim is to elucidate the philosophical underpinnings of the theory while contrasting it with the principal alternatives.

The first claim of the Geometrical Projection Theory is that every system of depiction determines a unique method of projection. Recall that *systems* of depiction are interpretive conventions which establish standards of accuracy for pictures. Methods of *projection*, on the other hand, are geometrical techniques for mapping three-dimensional scenes onto two-dimensional surfaces. According to the Geometrical Projection Theory the two are directly correlated. For example, the system of perspective line drawing uniquely specifies the method of projection for perspective line drawings. Each system encodes a distinct and stable projective rule of interpretation, its use in communication established by convention.[62] To understand a system of depiction, on this view, is to be able to apply its characteristic method of projection. Such an ability need not be based on explicit knowledge. Just as language users rely on tacit knowledge of abstract grammar rules to guide interpretation, successful pictorial communication requires only a tacit knowledge of methods of geometrical projection.[63]

[62] Jakobson (1987 [1921], 21) again: "The methods of projecting a three-dimensional space onto a flat surface are established by convention; the use of color, the abstracting, the simplification, of the object depicted, and the choice of reproducible features are all established by convention."

[63] See Hagen (1986, 84-5). Admittedly, in the typical case, competent *expression* in a system of depiction requires formal

The second component of the Geometrical Projection Theory specifies the role of geometrical projection in any system's definition of accuracy: a picture bears the relation of accurate depiction *to* a scene if and only if it is possible to derive the picture *from* the scene by geometrical projection. For an artist to intend to draw accurately is for the artist to intend (perhaps unconsciously) to perfectly execute a geometrical projection of some intended scene. Note that this is only an *in principle* projective mapping from pictures to scenes, not a *physical* process by which pictures are generated, nor a *computational* process by which projections are calculated. Like all descriptive semantic theories, the Geometrical Projection Theory articulates the conventional standards which govern the creation and interpretation of pictures, but imposes no constraints on how creation and interpretation are actually carried out.

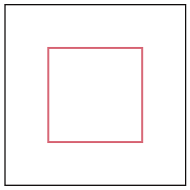
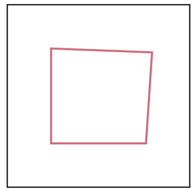
It follows from this definition of the Geometrical Projection Theory that any pictorial semantics based on it, though broadly of a kind with semantics for *languages* will differ fundamentally in underlying structure. Pictorial representation is sometimes described as "natural", in analogy with the phenomena of "natural meaning" exhibited by shadows, tree rings, or foots prints.[64] On the view of depiction as projective transformation, defended later, accurate pictorial representation falls between linguistic meaning and natural meaning in a precise and intuitive way. Like linguistic meaning, depiction is conventional; there are many possible systems of depiction which must be selected amongst for successful communication. But like natural meaning, and unlike language, depiction is not mediated by arbitrary associations of sign and content. Instead, depiction and natural meaning alike convey information in a more direct and unmediated way, in virtue of general, rule-based transformations. In the case of natural meaning these rules are the causal laws of physics. The rules underlying systems of depiction are the abstract laws of geometrical projection.

The best positive evidence for the Geometrical Projection Theory is simple induction over cases. Suppose two artists set out to draw a cube S from a given vantage point, relative to the system of perspective line drawing. The first picture, (5), is created in such a way that it conforms with perspective projection. And it seems that (5) is an accurate pictorial representation of S in the system of perspective line drawing. But the second picture (6), even though it deviates only slightly from (5), is such that there is no way that (6) could be derived from a cube by perspective projection. (The validity of this claim will become apparent in the next section.) Furthermore, (6) is an *in*accurate pictorial representation of S in the system of perspective line drawing. Thus accuracy in a perspective system and perspective projection appear to covary precisely, exactly as the Geometrical Projection Theory requires. This thought experiment may be carried out indefinitely on any given scene built from simple geometrical solids, and multiplied indefinitely with appropriate modifications for any other system of depiction.[65]

training, while *comprehension* does not. This is rather different from the case of language, where mastery of expression and comprehension are regularly achieved without formal training. But all this shows is that implicit knowledge of the rules of pictorial interpretation does not automatically entail an ability to draw. And it should also be emphasized that even formal artistic training never takes the form of fully precise and explicit drawing rules, but rather of loose heuristics the likes of which require a prior understanding of the relevant standards of accuracy.

[64] The term "natural meaning" is due to Grice (1957, 378).

[65] In particular, the evaluation can be carried out for the system of perspective line drawing with any given scene built from polyhedra— geometrical solids with only flat faces. The restriction is necessary because, in the crude method of

(5)  (6) 

A pictorial semantics ultimately determines a mapping from pictures to content. At the outset I defined the content of a picture as the set of scenes it accurately depicts. According to the Geometrical Projection Theory, a picture accurately depicts a scene just in case it can be derived from the scene by geometrical projection. Putting these definitions together yields the following analysis of pictorial content:[66]

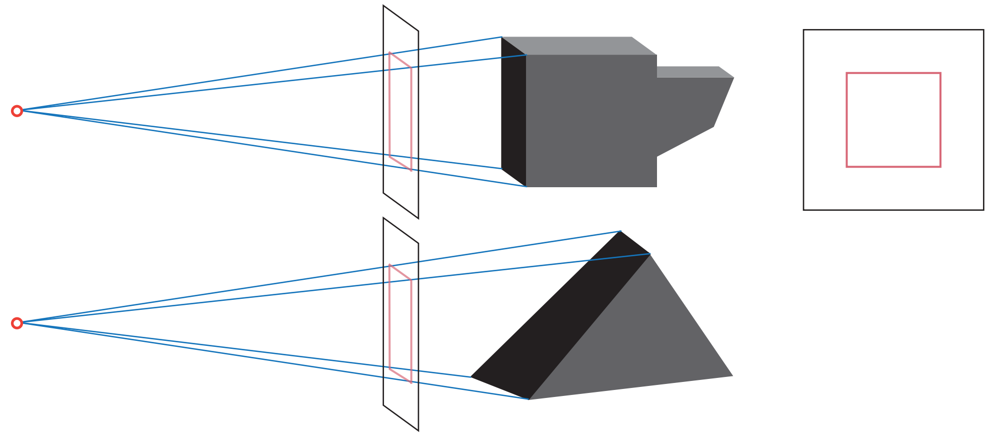
(7) ***Geometrical Projection Theory for Content***
The content of a picture in a given system is the set of scenes from which it may be derived by the projective method characteristic of that system.

Thus, to grasp what a picture depicts is to understand the space of possible scenes from which it could be projected. Those features held in common among the scenes in a picture's content correspond to how it represents the world as being; those features which vary among the scenes in a picture's content correspond to features of the worlds about which the picture is non-commital.

So, for example, the content of the drawing at right below will include all and only scenes containing shapes with visible faces which I will loosely describe as "square-like", because these are the scenes the picture could have been projected from. This "square-like" property is the feature which the picture represents the world as instantiating. But the picture does *not* represent the world as containing a *cube* or even a shape with a *square* face necessarily. For there are infinitely many irregular shapes, like those illustrated below, whose irregularities are not preserved in the given perspective projection.

edge-to-line mapping defined above, curved surfaces fail to register on the picture plane. Later we will consider more sophisticated projective methods which correct this problem. In the next section, we will also see that the claim of precise covariation requires some qualification.

[66]The analysis of content provided here is superficially like that proposed by Kulvicki (2006, p.57-9). But Kulvicki and I reach these conclusions by very different routes. Kulvicki holds that content of a picture is (or is defined by) its projective invariants (but offers no explicit account of depiction)— a view I argue against specifically in Greenberg (2010). I hold instead that content of a picture is its accuracy conditions, and define accurate depiction in terms of projection. These conceptual differences result in subtly but crucially divergent precisifications of the formulation here. See footnote 93 for discussion. In addition, Kulvicki holds that perspective projection is, in some sense, the prototypical method of projection, while I believe nothing of the sort.



Furthermore, because geometrical projection in general is insensitive to the identity and typological properties of the scene projected, such features are also left out of a picture's content. For example, the image of Obama presented earlier could equally well be a projection of Obama, a wax-sculpture of Obama, or suitably arranged lines in the sand. Thus its content would not differentiate between these alternatives.[67] As a consequence, a purely projective semantics like the one offered here is only suitable as a theory of spatial and chromatic *explicit* content. It is not intended to capture the rich, typological *extended* content sometimes ascribed to pictures (e.g. that this is a picture *of Obama*, or *of a man*).[68]

It might seem incredible that a semantics whose output is content as abstract and unspecific as this could play any role in an explanation of how novel pictures are quickly and reliably converted into useful information. But it is important to remember that pictures, like any kind of sign, are always interpreted against a backdrop of world knowledge. The same content which is quite unspecific taken as a whole, also contains quite specific *conditional* information that is elicited under the right presuppositions. A semantics based on geometrical projection provides the systematic and conventional rules which are necessary to extract bare-bones representations of the world from marked, flat surfaces. In concert with appropriate background knowledge about the world— as opposed to further knowledge of interpretive conventions— these admittedly skeletal representations can yield impressively precise and useful information.

For example, we saw how the content of the square drawing above is indeterminate between an infinite space of radically different objects. But suppose that we are independently informed that this picture accurately depicts a scene containing exactly one Platonic solid. (Roughly, a three-dimensional solid is Platonic if all of its faces are regular polygons of the same kind.) On its own, this knowledge restricts the space of possible scenes to *five*, corresponding to the five Platonic solids, but not fewer. Yet the square picture could only be a projection of *one* Platonic solid, the cube. Thus the content of the picture and background knowledge together entail an exact description of the scene in question— information which is more specific than either source on its own.[69] In the same

[67]Similarly, I understand "ambiguous" images like the Necker cube or duck-rabbit illustration to be merely semantically indefinite between, in the latter case, duck-scenes, rabbit-scenes, and infinitely many others. For psychological reasons anterior to the semantics, the duck and rabbit interpretations are especially vivid to human viewers.

[68]See §1.2 for further discussion of this distinction, and footnote 30 for its sources in the literature.

[69]Of course, the background knowledge in question need not put direct constraints on the content of an image in order

way, the Obama drawing considered in isolation could equally well depict lines in the sand, wax-sculptures, and infinitely many other oddities. But once we are informed that it accurately depicts a *human*, and given our background knowledge about humans, we are thereby provided with a mass of quite precise information about the physical characteristics of the person in question. When such supplementary correspondences between picture and world are specifically intended by the artist, it may be appropriate to construct a model of *rich* pictorial content, where these mappings are built into the content itself. But this development lies beyond the scope of the present inquiry.[70]

2.3 Alternative theories

The Geometrical Projection Theory takes its inspiration from the study of projective geometry and recent work in computational perceptual science.[71] But the view differs sharply from the theories of depiction which have dominated contemporary Philosophy of Art, each of which has evolved from a correspondingly different point of origin.[72] In what follows I distinguish the Geometrical Projection Theory from the three main proposals discussed in this literature. Here two provisos are necessary. First, I do not assume that these theories are mutually exclusive; they represent distinct areas of emphasis in a spectrum of actual and possible proposals. Second, I will treat all the theories under discussion as theories of *accurate* depiction, though in the literature they are typically stated simply as theories of *depiction*. I hope the translation does not substantively distort the original intent. Given the constraints of space, I am not able to pursue detailed objections and replies; my aim is merely to distinguish The Geometrical Projection Theory from the alternatives.

(i) The *Symbol Theory* of depiction articulated by Nelson Goodman (1968, 5) begins with the recognition that communication with pictures, as with language, is governed by conventional systems of interpretation. Impressed by this analogy, Goodman went on to hold that depiction is based on a library of arbitrary correspondences between pictures and their content, much like the lexicons of language.[73] Yet this view is implausible on its face: the relationship between, say, a drawing of

to allow a viewer to extract rich content from it. Consider an agent who has antecedent knowledge that a picture accurately depicts some actual scene— because of her knowledge of the reliability of the artist— but who is otherwise wholly ignorant about what it depicts. Given background knowledge about what *the world* contains, she may come, after viewing the picture, to conclude that it accurately depicts a cube, because there is nothing else in the world it could accurately depict. Richer interactions, which make way for more measured confidence in the reliability of the artist and the furniture of the world, are the characteristic domain of Bayesian models of interpretation.

[70] See Greenberg (2011) for further discussion.

[71] An active research program in computational perceptual science over the last forty years has attempted to articulate ever more complete algorithms for the interpretation of line drawings. (Early contributions include Huffman (1971), Clowes (1971), and Waltz (1975). Two excellent overviews of this field are Willats (1997, ch. 5) and Durand (2002).) The principles evolved in this literature are not general standards of correct interpretation. Rather, they are implementable algorithms, based on reasonable heuristic assumptions, for determining reasonable interpretations. (e.g. Huffman 1971, 297-9) Pictorial semantics occupies a distinct but compatible level of theoretical description: it articulates an abstract mapping from pictures to information which computational theories of interpretation aim to make practically computable. (See Feldman (1999) for a discussion of the analogous distinction for vision.) The present work owes a special debt to the discussion of depiction in Willats (1997).

[72] Among proposals by philosophers, the account of Hyman (2006, ch. 6) most nearly prefigures the Geometrical Projection Theory. Hyman's account is based on the recognition that projection is a necessary condition for depiction. But the similarities with the present view largely end there. For example, Hyman does not construe his account as a theory of accuracy; his proposal is highly informal, with little detail about the role of the projective index, or the connection between the geometry of projection and the definition of pictorial line and color (see §3 for my own position on these matters); and he seems to assume that perspective projection in particular is central to the definition of depiction, which I reject.

[73] Goodman would have objected to the label "symbol theory", but not the characterization in terms of arbitrary mean-

Obama and the scene it accurately depicts is much more intimate and direct than the admittedly arbitrary connection between, say, the word "tree" and the property of being a tree (or a given tree). Furthermore, as Wollheim (1987, 77) has pointed out, the theory fails to explain how we are able to interpret novel images with such facility, for there is no straightforward way of extrapolating from a list of arbitrary associations to new cases.

According to the Geometrical Projection Theory, by contrast, accurate depiction is based on *general* rules of projective transformation, not a long list of arbitrary meaning associations. On this view, the interpretation of novel images is straightforwardly explained as the application of a familiar general rule to unfamiliar particular cases. To determine whether a novel image accurately depicts a given scene, one must simply consider whether the former can be derived from the latter by geometrical projection. Meanwhile, the apparently intimate relationship between an accurate depiction and its subject is explained by the relative simplicity of such projection rules and the speed with which they are applied— aided no doubt by the computationally powerful visual system. Of course, the projection analysis is compatible with the fact that systems of depiction, like languages, define conventional rules of interpretation. But the theory rejects the conflation of these rules with the arbitrary lexicons of language. Instead, each system is a convention for interpreting pictures according to a one of the many possible methods of projection.

(ii) *Resemblance Theory* is the oldest and most orthodox view of depiction.[74] Its starting place is the powerful intuition that a picture, say of Obama, and the scene it accurately depicts are substantively similar with respect to certain color and shape properties. This observation has inspired the following view, roughly stated: a picture accurately depicts a scene just in case the picture and the scene are similar in the appropriate respects (and, in many accounts, the artist intends the picture to represent that scene).[75] Resemblance theorists have widely acknowledged the importance of geometrical projection in specifying the right relationship between pictures and scenes, but have attempted, sometimes with considerable effort, to fold this fact into a genuinely similarity-based analysis.[76]

The Resemblance Theory has been subjected to a host of objections, each met with ever more sophisticated revisions.[77] Yet as I argue in Greenberg (2010), the very structure of the theory renders it empirically inadequate. The basic point can be made with respect to the system of CURVILINEAR PERSPECTIVE, as in the "fish eye" photograph below. Due to its characteristic curvature, the image clearly does *not* resemble its subject in the intuitive and direct way exhibited by standard perspective depictions. Resemblance theorists have tried to preempt this concern by claiming that such curvature is semantically inert, inconsequential to the content of the image, which they insist on

ings. Lopes (1996, 65) notes that Goodman also resisted the view that systems of depiction are conventional in any obvious sense. Bach (1970) is among the few adherents of Goodman's view.

[74]In the 20th and 21st centuries, resemblance theorists have included Peirce (1906), Wittgenstein (1921/1997), Morris (1946), Gibson (1960, 1971), Manns (1971), Novitz (1975), Fodor (1975), Neander (1987), Peacocke (1987), Malinas (1991), Files (1996), Hopkins (1998), Newall (2003), Kulvicki (2006), Blumson (2007, 2009a), and Abell (2009). Resemblance has also found adherents in the adjacent domains of mental representation and scientific representation, typically under the banner of "isomorphism"; see Cummins (1996), French (2003), and Giere (2004, §4).

[75]Resemblance theorists vary considerably in how they revise or extend this rather crude formulation.

[76]For example: Peacocke 1987; Hopkins 1998, 71-81; and Kulvicki 2006, ch. 3-4.

[77]The most fundamental objections have been those posed by Descartes (2001 [1637], 90), Goodman (1968, 3-10, 34-9), and Lopes (1996, ch. 1). Recent, sophisticated forms of Resemblance Theory have been proposed by Kulvicki (2006), Blumson (2007) and Abell (2009).

analyzing in terms of similarity. Yet I showed that just the opposite is the case. The accuracy of a picture in curvilinear perspective depends on the precise degree of deviation between the straightness of edges in the scene and curvature of lines in the picture. In this and many other systems, accuracy depends as much on a picture's systematic *differences* with the scene depicted as on its similarities. It follows that these systems cannot be analyzed in terms of similarity alone. Instead, a more general notion of transformation is required; I propose geometrical projection.

Both the Geometrical Projection Theory and the Resemblance Theory reject the view that accurate depiction is based on arbitrary associations between pictures and scenes; and both hold that it is grounded instead in a spatial and chromatic relation of "fit" between a picture and its subject. But they differ fundamentally in their construal of this relation. Formally, the Resemblance Theory holds that the relation underlying accurate depiction is a *symmetrical* relation of similarity, but for the Geometrical Projection Theory it is an *asymmetrical* relation of transformation. These relations are conceptualized very differently in each case: whereas the Resemblance Theory claims that accurate depiction arises from the *commonalities* between the picture and the scene, the Geometrical Projection Theory holds that depiction arises from the fact that the picture may be *derived* from the scene by projection. This shift in conception buys the latter a theoretical simplicity and empirical validity that resemblance accounts fall short of.

At the same time, those apparent similarities between pictures and scenes which have seemed so powerful to so many theorists are not ruled out. For any kind of transformation, there are INVARIANTS— properties which are always held in common between the inputs and outputs of the transformation.[78] The transformational invariants for common methods of projection are well known. For example, under perspective projection, straight lines in a scene are always mapped onto straight lines in the picture; so straightness of line is an invariant. (The inverse does not hold generally.) A consequence is that "betweenness" relations are also invariant. By contrast, parallelism is *not* an invariant; in perspective, parallel lines are usually mapped onto converging lines. By understanding the alleged resemblances between pictures and scenes as transformational invariants, the intuition of similarity may be substantiated even while the proposal to ground depiction in resemblance is rejected.

(iii) Finally, *Perceptual Theories* follow from the striking analogies between the process of picture interpretation and that of normal visual perception; depiction is then analyzed in terms of the kinds of perceptual effects a picture has on suitable viewers. The crudest perceptual account holds that a picture accurately depicts a scene if and only if seeing it causes a viewer to have the *same* perceptual experience she would have upon seeing the scene.[79] Depending on the theory, more sophisticated

[78]See Hagen (1986, ch. 2).

[79]Such a view is defended by Gibson (1954), and often ascribed to Gombrich (1960), though it is unclear if the ascription is entirely apt.

variants hold that a picture accurately depicts a scene if and only if it permits a viewer to see the scene *in* the picture, it facilitates the viewer's *recognition* of the scene, or it is a suitably vivid prop with which the viewer may *pretend to see* the scene, and so on.[80]

Despite their diversity, all Perceptual Theories face the same basic problem of giving precise articulation to their preferred notion of a "suitable viewer". The problem arises because not just any viewer will fit the designated theoretical role. There is no hope, for example, of defining accurate depiction in terms of the perceptual reactions of a chronic hallucinator or an acute myopic. Instead, Perceptual Theories require that the perceptual system of a suitable viewer must be, in some sense, ideal. In order to give substance to her proposal, it is incumbent on the Perceptual Theorist to specify the operative notion of ideal perception. And because perceptual abilities vary so widely, is not clear that such an idealization can be derived straightforwardly from the generalizations of contemporary perceptual science, or from statistical regularities. That is not all: in order to capture the difference in accuracy conditions between various systems of depiction, a Perceptual Theory must posit differences in perception for each. But then we are owed an account not merely of *ideal* perception, but of *ideal black and white* perception, *ideal color* perception, *ideal line drawing* perception and so on. What are these different modes of perception? Perceptual Theorists have had little to say in the way of specific answers to this question. But without these details explicitly worked out, Perceptual Theories cannot claim to offer a substantive analysis of accurate depiction.

The Geometrical Projection Theory, by contrast, offers a significant advance over Perceptual Theories by explicitly and precisely defining the conditions of accurate depiction in terms of geometrical transformation. (In the next section and in the formal presentation of the Appendix, I show that this definition can be quite exact.) The many systems of depiction do not engender many modes of ideal perception, but merely many varieties of projection. Plausibly, if there were an ideal perceiver of pictures, her visual system would perforce follow the geometrical rules supplied by the Geometrical Projection Theory. But the theory need not posit any such creature; it characterizes the rules directly, without mediation. This is not to say that the forms of geometrical projection at work in common systems of depiction are divorced from visual perception. Instead, they are abstractions and strategic extensions of the *optical* projection of light into the eye. Thus the computations required to interpret a picture are broadly of a kind with those required of the visual system to interpret patterns of illumination reflected to the eye in normal perception.

Of course, Geometrical Projection Theory on its own does not explain why only a narrow range of projective methods are actually implemented in normal human pictorial communication. There are many possible systems of depiction, and some, like those that scramble their projection lines in eccentric ways, are never actually used. A plausible explanation is that although the human visual system can be trained to interpret a variety of projective methods that diverge from those that figure in normal perception, this capacity of limited. Ultimately, only a narrow range of projective methods can be computed by that part of the human visual system which is harnessed for the interpretation of pictures. Yet this appeal to visual biology is philosophically innocent. Whereas

[80]The "seeing-in" theory is due to Wollheim (1987, 46-59; 1998). Schier (1986, ch. 3) and Lopes (1996, ch. 7) defend recognition-based views. Walton (1990, ch. 8) is the originator the pretense theory. Loewer and Godbey (1978, 339-40) and Wollheim (1998, 217-8) succinctly capture the core commitments of this group of theorists.

Perceptual Theories rely on the workings of *ideal* viewers, the Geometrical Projection Theory appeals only the capacities of *actual* viewers to explain why only a narrow range of projective methods are *actually* used. Systems of depiction are geometrical abstractions, just as generative grammars are logical abstractions; the extent and variety of such abstractions which can be successfully computed by the organs of human cognition is matter of ongoing empirical inquiry.

This concludes my general presentation of the Geometrical Projection Theory. I've shown how the theory yields a pictorial semantics— a systematic and conventional mapping from pictures to content— and I've indicated why it realizes the aims of pictorial semantics better than the alternatives. If the theory can be stated explicitly and precisely for at least one system of depiction, then I will have begun to meet the principle challenge facing the Pictorial Semantics Hypothesis. In the next section, I show that such precision can be achieved.

3 Semantics for Perspective Line Drawing

Having laid the foundations of the Geometrical Projection Theory, I now turn to developing it in detail for the particular case of perspective line drawing.[81]

Swiddle@www.skyscrapercity.com

Though this is just one of infinitely many possible systems of depiction, it has witnessed an extraordinary rise in global popularity since its codification in the Italian Renaissance. Today, drawings made in perspective are the norm, and nearly all mass market camera lenses are designed to mimic the geometry of perspective depiction.[82] Its most notable characteristic is that more distant objects are represented by smaller elements in the image, like the railroad ties at right; consequently, receding parallel lines in the scene are represented by converging lines in the image.[83]

This system is too complex to analyze systematically in the space of just one paper. Instead, my method here will be to develop an exact semantic theory for a recognizable *fragment* of the system of perspective line drawing.[84] According to the Geometrical Projection Theory, all systems of depiction are defined in terms of their characteristic method of projection. So I will begin in §3.1 by describing in some detail the method of projection associated with this simplified version of perspective line drawing. §3.2 then integrates this account into a precise semantics for the system

[81] What I am calling "perspective" here is sometimes referred to as "linear perspective", in contrast with "curvilinear perspective" which involves projection onto a curved projection surface.

[82] More exactly, mass market camera lenses are designed to mimic *linear*, as opposed to *curvilinear* perspective; they yield merely perspectival representations more or less by default. See Pirenne (1970) for discussion.

[83] Here it is crucial to observe a distinction between the geometry of perspective *depiction* and the geometry of human visual *perception*. Although perspective was developed over many centuries by artists and scholars attempting to recreate perceptual experience on paper, the one is ultimately a product of convention, the other biology. (See: Da Vinci et al. 1970 [c. 1500], II.107-109; Gombrich 1960, 4-30; Gibson 1978, 232; Hyman 1995.) In fact, recent empirical research by Rogers and Rogers (2009) strongly suggests that some roughly perspectival but *non*-linear system best models the structure of visual perception. Many have been led to the same conclusion by less quantitative methods: Helmholtz 1962 [1867]; Pirenne 1970; Hansen 1973; Arnheim 1974; and Hansen and Ward 1977.

[84] This is the approach pioneered by Montague (1970a, 1970b) in the study of natural language semantics. See Partee (1980, 1-5) for discussion.

of depiction in question. An explicit formalization of this analysis is included in the Appendix. Finally, §3.3 considers how the semantics for this fragment might be extended to match the systems of line drawing found in ordinary pictorial communication.

3.1 The method of perspective projection

A method of projection is a kind of algorithm for mapping three-dimensional scenes onto a two-dimensional picture planes.[85] Following Willats (1997), this method may be divided into two stages. In the first stage, an array of connecting lines are established, linking points in the scene to point in the picture plane. In the second, properties of the scene are mapped onto colors in the picture plane, according to a spatial distribution determined by the connecting lines of the first stage. Methods of projection can be defined by how they carry out these two stages: the first stage is specified by a given method's PROJECTION CONDITION, the second by its DENOTATION CONDITION.[86]

In any method of projection, we begin by positioning a finite PICTURE PLANE in relation to a given scene, and then defining a set of PROJECTION LINES which establishes a system of correspondences between points in the picture and points in the scene— illustrated below.[87] For every point in the picture plane, exactly one projection line passes through it. The projection condition structures this set of correspondences by specifying a certain configuration of projection lines. The projection condition characteristic to methods of *perspective* projection requires that all projection lines converge on a single point, commonly called the VIEWPOINT. This ultimately determines the "point of view" of the picture.[88]

Here it is often convenient to speak of the PROJECTIVE COUNTERPARTS of a point on the picture plane: these are the points in the scene to which it is connected by a projection line.[89] Typically, not *all* projective counterparts of a point in the picture are relevant to the final image. Only those that are ACCESSIBLE from the viewpoint in the following sense are of consequence: they can be reached by a projection line without passing *through* a surface in the scene. Projection lines from picture points to their accessible counterparts can be thought of as "lines of sight", since, like sight, they link the viewpoint to those surfaces in the scene which are not occluded by any other surface in the scene. For visual clarity, I have indicated only those projection lines below which connect the viewpoint to the accessible corners of the object in the scene.

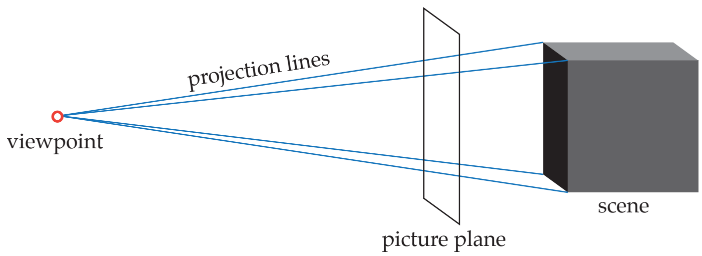
[85] The mathematical details of the following discussion are based on presentations in Sedgwick (1980), Hagen (1986, ch. 2-5), and Willats (1997, ch. 2, 5).

[86] Willats (1997) correspondingly organizes methods of projection into "projection systems" and "denotation systems". See Durand (2002) for an extension and critical discussion of this approach.

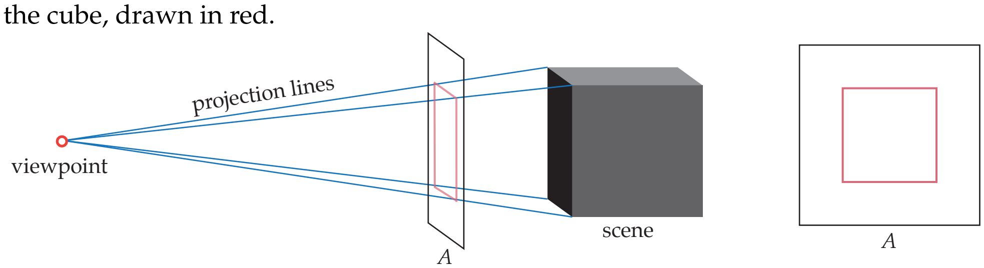
[87] Admittedly, "picture plane" is a misnomer; I mean it to refer to a contiguous finite region of a plane.

[88] The viewpoint is sometimes called the "station point"; I will also refer to it as the "vantage point." Note that the viewpoint is not the same thing as a picture's "vanishing point". The viewpoint is a position in the space of the scene, outside the picture plane, introduced to define a perspective projection. A vanishing point is any position in the space of the picture plane to which two lines depicting parallel lines in the scene will converge. There may be as many vanishing points in a picture as there are such pairs of lines. The distinction between "one-point", "two-point", and "three-point" perspective describes vanishing points, not viewpoints.

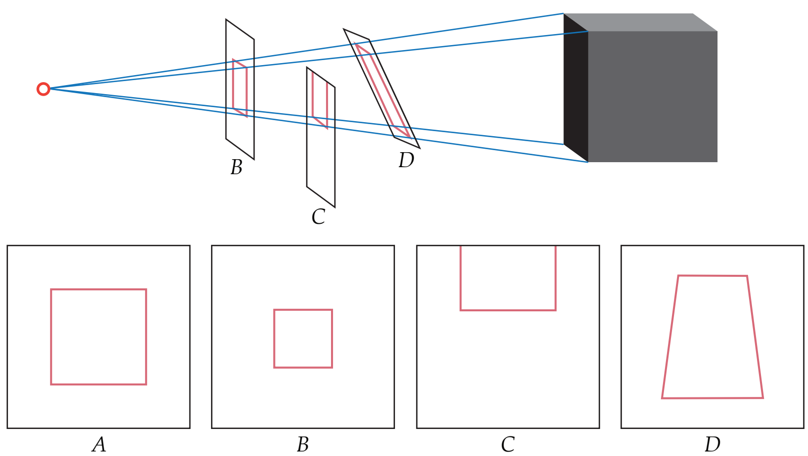
[89] Some points in the picture, for example at the periphery, may not have projective counterparts in the scene.



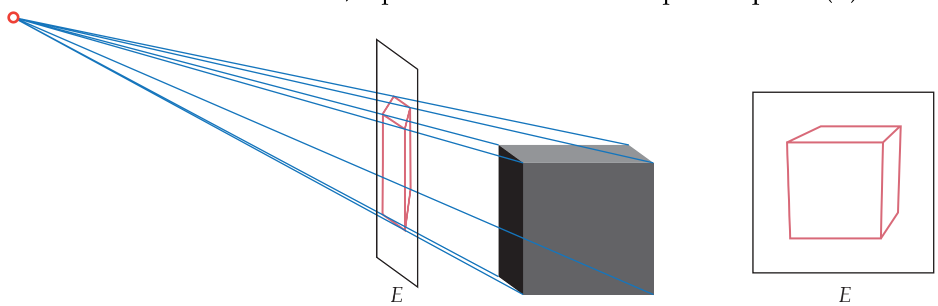
The next stage in the projection is specified by the denotation condition— the rule by which features of the scene are mapped onto features of the picture plane along the lines of projection. In this case I'll use a simplified version of the rule characteristic of *line drawing*: a point in the picture is colored red (for example) just in case it has a projective counterpart that lies on an accessible edge in the scene. I'll define an EDGE, for now, as the intersection of two flat planes. This rule is applied to every point on the picture plane; those points whose color is thereby left undefined assume the background color of the "page." If the picture plane is now seen face on, as at right, it reveals a side view of the cube, drawn in red.



We can further manipulate the relative positions of the picture plane and viewpoint, with predictable consequences for the resulting image. The following figure illustrates the results of altering the location and orientation of the picture plane. Shifting the plane closer to or farther from the viewpoint (*B*) has the effect of altering the scale of the resulting image. Shifting the picture plane vertically (*C*) causes the projection of the scene to drift from the center. When the picture plane is tilted (*D*), the resulting image records exactly the same aspects of the cube as (*A*), but introduces the characteristic "railroad-track effect" of perspective projection, where edges of the cube which are in fact parallel are now represented by converging lines.



Each of the alterations just described changes the *way* the scene is represented, but none reveal additional information about the scene not already reflected in (A). By contrast, when the viewpoint is shifted, new features of the scene are revealed. For example, let us move the viewpoint above and to the side of the cube, while simultaneously sliding the picture plane to intercept the projection lines. The principle effect of repositioning the viewpoint in this way is that it can now "see" two additional faces of the cube, represented below on the picture plane (E).



Images (A)-(E) are produced by the same general method of projection, their differences owing to different selections of picture plane orientation and viewpoint. In general, there is no such thing as "the perspective projection" of a scene, independent of such parameters. But *relative* to a choice of positions for picture plane and viewpoint, the method of perspective projection delivers a unique projection of any scene. Here it is convenient to collect picture plane and viewpoint position together into a single PROJECTIVE INDEX. Then we may say that, relative to a projective index, perspective projection determines a unique projected image of any scene.

Let me summarize. A method of projection is an algorithm for mapping a three-dimensional scene onto a two-dimensional picture plane. All methods of projection are defined in terms of an array of projection lines where exactly one passes through each point on the picture plane. The method of projection for perspective line drawing can then be identified first by how it structures these projection lines, and second by how it uses them to determine a mapping of features from scene to picture, its projection and denotation conditions respectively:

Projection: Every projection line intersects the viewpoint.

Denotation: A point in the picture is red if and only if it has a projective counterpart in the scene which lies on an accessible edge.

Together, these conditions define a method of projection such that, given a scene and a projective index— that is, the position of the viewpoint and picture plane— the method determines a unique projected image. Thus where S is a scene and j a projective index, we may speak of *the* perspective projection of S relative to j. An explicit formalization of the method of perspective projection is provided in the Appendix.

Other methods of projection can be derived by independently modifying the projection and denotation conditions. For example, methods of *parallel* projection differ with respect to the projection condition: the viewpoint of perspective projection is replaced by a view *plane*, such that all projection lines intersect it at right angles. On the other hand, methods of *color* projection differ with respect to the denotation condition: instead of correlating red points in the picture with edges in the scene, a rudimentary color system might hold that points in the picture picture simply have the same color as their accessible projective counterparts in the scene. And so on. Formal definitions of these alternatives are also included in §**??**. For now, I turn to integrating the method of perspective projection into a theory of accuracy.

3.2 Semantics for the system of perspective line drawing

The method of perspective *projection* just illustrated is an algorithm for deriving pictures from scenes. The *system* of perspective line drawing, on the other hand, determines a standard of pictorial accuracy. I shall say that every possible composition of red lines on a bounded white plane corresponds to a picture belonging to the system of perspective line drawing.[90] For any such picture and any scene, a definition of accuracy in this system determines whether the picture accurately depicts the scene. According to the Geometrical Projection Theory, the system of perspective line drawing defines accuracy in terms of perspective projection in roughly the following way. For any picture P in the system and scene S:

(8) P accurately depicts S in the system of perspective line drawing if and only if
P is a perspective line drawing projection of S.

To state this formula more precisely, let L be the system of perspective line drawing suggested by the projective method above. Where a picture P accurately depicts a scene S in system L, we may write: "P accurately depicts$_L$ S". Now, since the specified method of perspective projection was entirely determinate, it can be represented by a mathematical function, what I will call a PROJECTION FUNCTION. A projection function takes as inputs (i) a scene and (ii) a projective index, and outputs an image— the unique perspective projection of the scene relative to the index. Where $\mathsf{proj}(\cdot)$ is the perspective projection function, S a scene, and j a projective index I shall say that $\mathsf{proj}(\cdot)$ applied to S and j returns an inscribed picture plane P, that is, $\mathsf{proj}(S, j) = P$. Then a first attempt to render (8) in more exact terms may be articulated. For any picture P in L and scene S:

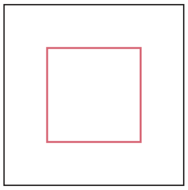
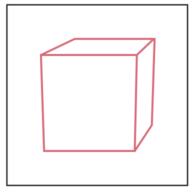
(9) P accurately depicts$_L$ S iff $\mathsf{proj}(S, j) = P$.

That is: P accurately depicts S in L if and only if the perspective projection of S relative to j is P. The problem with this definition is that it makes the projective index j part of the system itself. This would mean that all accurate perspective images would have to be drawn from the same viewpoint. But this is clearly false: both (11) and (12) below are accurate perspective depictions of S, but from different viewpoints. The solution is to allow some variability in j. One way to achieve this is to existentially quantify over the projective index on the right-hand side of the equation.

(10) P accurately depicts$_L$ S iff there is some j such that $\mathsf{proj}(S, j) = P$.

On this analysis, two distinct projections of a scene, such as the two views of the cube below, are both accurate representations of that scene, since both can be projected from that scene according to *some* index.

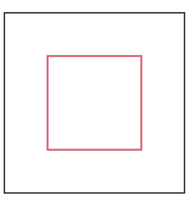
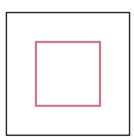
[90] This does not imply that every marking of red ink on paper specifies a picture in this system, just as merely sounding like a word of English is not sufficient for being a word of English. The sign must be created with the proper intent under the appropriate conditions.

(11)  (12) 

I concede that there is a strand in the concept of accurate depiction that conforms with this analysis, but my interest here is in a more specific notion. According to this alternative, accurate depiction only obtains in relation to a projective index, and artists always intend that their pictures be interpreted relative to a particular index. If the picture is not a perspective projection of the scene from that index, then it is not accurate — *even if* it is a correct projection of the scene from some other index. Thus, if two artists create (11) and (12) respectively, but both intend to represent the scene from the same vantage point, then only one succeeds in accurately depicting the scene. To accommodate this idea, I now analyze depiction as a relation between a picture and a scene relative not only to a system of depiction, but *also* to a projective index.[91] Where P accurately depicts$_L$ relative to index j, I will write "P accurately depicts$_L$ at j". Then for any picture P in L, scene S, and index j:

(13) P accurately depicts$_L$ S at j iff $\mathsf{proj}(S, j) = P$.

Unfortunately, this implementation of the Geometrical Projection Theory is too demanding in a subtle but important way. It is not the case that an accurate picture can always be derived by projection from its subject; sometimes a picture is accurate when it is merely *similar* to some other picture which *can* be so projected. This principle is plainly illustrated by considerations of scale. Suppose that two artists intend to draw a cube from a given vantage point, and produce (14) and (15) below. Both images are perfectly accurate representations of S, despite their difference in size. Yet this flexibility is not allowed by (13), for $\mathsf{proj}(S, j)$ yields a picture plane of fixed size, and any deviation from this picture leads to a failure of the biconditional. To accommodate this, (13) should be revised; rather than requiring that the picture itself be derived from the scene by projection, I instead require that some *copy* at the same or different scale be so derived.

(14)  (15) 

Other systems of depiction require other dimensions of freedom besides scale, including flexibility with respect to choice of color, line smoothness, and in extreme cases, overall metrical structure.[92] We need a general way of handling these various cases. To do this, let us introduce a

[91] See Hyman (2006, 82).

[92] Some color systems tolerate constrained variations in color among accurate depictions of the same scene. (Thanks to Dom Lopes for bringing this to my attention.) Some line drawing systems tolerate limited deviations from perfectly straight lines in the depiction of straight edges. Subway maps exhibit constrained deviations from the metrical structure of a strict projection, while preserving quasi-topographic features like station-betweenness. (Willats (1997, 70-7); Giere 2004, §4). An alternative approach is rather natural at least the case of wobbly lines. Pictures, considered as abstract objects, may differ in interesting ways from the marks which express them— just as sentence structure can depart from the "surface structure" of an utterance. Then we might think that it is only marks on the page which are wobbly; they express pictures composed of perfectly smooth lines. I have no particular complaint with this approach, but presumably it should also be extended

FLEXIBILITY FUNCTION, $\mathsf{flex}(\cdot)$, which takes a picture and returns the set of suitably related pictures for that system of depiction. Now, instead of requiring for accuracy that the projection of S be *identical* with P, we merely require that the projection of S be *within* $\mathsf{flex}(P)$, that is, the class of pictures appropriately similar to P. On this approach, the flexibility and projection functions play against each other, for while the projection function constrains the conditions on accurate depiction, the flexibility function weakens these constraints. But note that the idea here is not to introduce an open-ended parameter; rather, we must define specific flexibility functions for for specific systems. Thus, for L, we might define $\mathsf{flex}(\cdot)$ as follows, for any picture P in L:

(16) $\mathsf{flex}(P)$ = the set of pictures which are scale variants of P.

With this amendment, we arrive at the final analysis:

(17) ***Accuracy in Perspective Line Drawing***
For any picture P in L, scene S, and index j:
P accurately depicts$_L$ S at j iff $\mathsf{proj}(S, j)$ is in $\mathsf{flex}(P)$.

This is a definition of accuracy for the system of perspective line drawing. It follows from this analysis plus my initial definition of depictive content that the content of a picture in perspective line drawing is, setting aside the flexibility function, the set of scenes from which it can be derived by perspective projection. To make this statement precise, one must account for the observation that accurate depiction holds not only relative to a system of depiction, but also to a choice of projective index. A natural view here is that pictorial content is defined relative to such an index as well.[93] Then just as linguistic content is determined in part by isolated features of conversational context, pictorial content depends on communicative context to specify an intended projective index. Formally, the content of a picture P in system L relative projective index j is, roughly, the set of scenes P could be derived from by perspective projection relative to j. More exactly:

(18) ***Content in Perspective Line Drawing***
For any picture P in L, scene S, and index j:
the content of P in L relative to j = the set of scenes S such that: $\mathsf{proj}(S, j)$ is in $\mathsf{flex}(P)$.

I have thus demonstrated, for one simple system of depiction, how the Geometrical Projection Theory determines a pictorial semantics— a systematic and conventional mapping from pictures to content. The analysis may be extended to other systems of depiction, and so to depiction in general, by defining alternative projection and flexibility functions. Once again, the details of such

to scale and color. But then pictures are so abstract— lacking determinate size, for example— that they become difficult to discuss without excessive technicality.

[93]This is not the only way. We might also define the content of P in L, independent of a projective index as the set of scene-index pairs $\langle S, j\rangle$ such that $\mathsf{proj}(S, j)$ is in $\mathsf{flex}(P)$. The difference between this view and the one expressed in the text is analogous to the difference between eternalist and temporalist theories of linguistic content. (Richard 1981) In each case there is a question about whether content is fixed by the value of given contextual parameter, or whether it generalizes over the possible values of this parameter. Kulvicki (2006, 57-9) holds yet another view, which derives from his assumption that content is determined by projective invariants; he defines the content of a picture, independent of projective index, as the set of scenes it could be projected from according to *some* index. (The resulting set is the set of scenes which share projective invariants with the picture itself.) But at least for the case of linear perspective, this proposal seem to me too crude: it misses the fact that by gaining information about a picture's viewpoint, we are able to extract more specific information from the content of the picture. Both the position in the text and that sketched above can accommodate this observation. For the case of curvilinear perspective, the view faces even more fundamental problems, as I argue in Greenberg (2010).

definitions are included in the Appendix.

3.3 Natural systems of perspective line drawing

By demonstrating that it is possible to state a precise semantics for one simple system of line drawing, I have shown that unqualified skepticism about the Pictorial Semantics Hypothesis is unwarranted. But even this may not satisfy the entrenched doubter: perhaps artificially constructed systems of depiction like L have semantics, she admits, but the prospects seem dim for the unregimented depictive styles actually found in ordinary communication. The aim of this section is to offer a partial response to such a challenge. First I'll highlight the sense in which the skeptic is right: by design, L is only a fragment of any natural system of line drawing. But then I'll show how recent developments in computational perceptual science can be used to expand this fragment in promising ways. These findings give us reason to hope that, for all their complexity, even natural forms of depiction are undergirded by systematic rules, capable of mathematically precise articulation.

Following a precedent of linguistic theory, we can roughly distinguish NATURAL from ARTIFICIAL systems of depiction. Natural systems have evolved organically in ordinary communication, while artificial systems are introduced and disseminated by solely by explicit definition. For the scientist of human behavior, natural systems are the primary quarry, but because of their stipulative origins, the semantics of artificial systems are easier to define. For this reason, well-understood but artificial systems like L serve as useful stepping stones on the path toward the analysis of the not-yet-understood natural systems. Thus, while L is recognizable as a system of perspective line drawing, it is only a fragment of any natural drawing system, for it is suitable for interpreting only depictions of geometrical solids with flat sides.

The chief shortcoming of L lies in the denotation condition for the method of perspective projection upon which L is based, reproduced here:

(19) A point in the picture is red if and only if it has a projective counterpart in the scene which lies on an accessible edge.

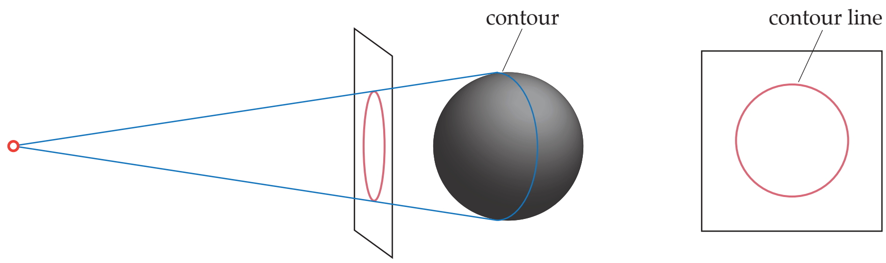
Recall that an EDGE was defined as the intersection of two flat surfaces. As a consequence, L cannot be used to depict curved objects like spheres, since spheres contain no intersecting flat surfaces. It doesn't help to expand the definition of edge to include what are sometimes called "creases"— abrupt intersections of any kind of surfaces— because spheres contain no such discontinuities. Indeed, given that spheres are completely uniform, it may seem puzzling how features of a sphere could be mapped to lines in a picture at all.

The answer is that we must introduce a very different concept of edge— one which applies not to intrinsic features of a scene, but instead to relational properties defined relative to the viewpoint.[94] Such is the notion of CONTOUR. Intuitively, a contour is any visible perimeter in a scene which occludes features of a scene behind it. Formally, we can define a contour as that region of a surface which is tangent to a projection line extended from the viewpoint— that is, the region of

[94] The idea of a "view-dependent" edge is described by Durand (2002, §6.2). Philosophical writing on depiction has employed the related notion of "occlusion shape", or its close kin (Peacocke 1987; Hopkins 1998, 53-7; Hyman 2006, 75-9).

a surface which a projection line "barely touches", as illustrated below. With this, we can define the denotation condition for a revision of L in such a way that yields successful renderings of a sphere. (Here a CONTOUR LINE is just a line in the picture which corresponds to a contour in the scene.)

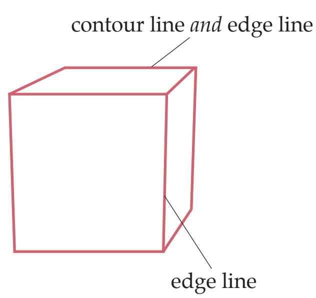
(20) A point in the picture is red if and only if it has a projective counterpart in the scene which lies on an accessible contour.



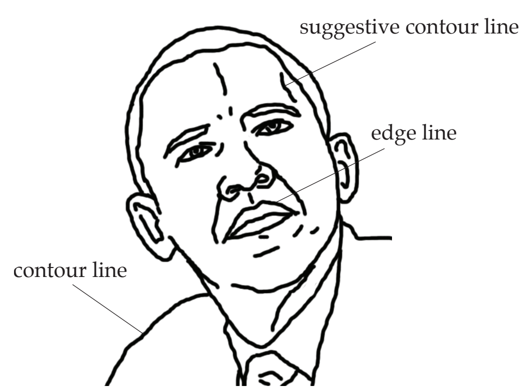
Yet contours cannot replace edges in the analysis of line, for while (20) would successfully capture the *silhouette* of a cube, it would fail to depict any of the cube's *internal* edges— those edges facing the viewpoint which occlude nothing from view. Instead, we must combine the two definitions, allowing that a point is inscribed on the picture plane when it has a projective counterpart in the scene which lies *either* on an accessible edge *or* an accessible contour. This definition yields successful depictions both of spheres, like that above, and of cubes, as in (21).

But even this innovation will not suffice. Common line drawings in perspective contain certain kinds of line that correspond neither to edges in the scene, nor to contours. Instead, they merely *suggest* the shape of the object depicted. DeCarlo et al. (2003) dub these SUGGESTIVE CONTOURS. Consider, for example, the suggestive contour line tagged in (22). It does not indicate the sharp intersection of two surfaces (an edge) nor is it naturally interpreted as describing a jutting cranial ridge hiding a valley behind it (a contour). Instead it merely suggests a certain structural arc at a certain location on Obama's head.

(21)



(22)



One might think that we have here finally reached the limits of systematic inquiry, that line drawing is ultimately organic and improvised, and that a semantic analysis will never be able to account for such expressive fluctuations as suggestive contour. But this skepticism is premature. In a recent paper from the field of computer graphics, DeCarlo et al. (2003) propose an ingenious analysis of suggestive contour lines. They offer an algorithm according to which suggestive contour lines are inscribed not when they correspond to *actual* contours relative to the viewpoint, but when they *would* correspond to contours under constrained *permutations* of the viewpoint. On this account, the presence of the suggestive contour line in (22) indicates that the shape of Obama's head at that point is such that, if the viewpoint were shifted slightly to the left, that region would emerge as a genuine contour. Such "nearby" contours obviously reflect structurally significant features of a scene; the suggestion is that, in the system of perspective line drawing, the rules for rending contours are naturally extended to capture this additional structure.

These considerations are only a small indication of the challenges facing the analysis of natural line drawing. The moral is that there remain substantive matters to work out in the elaboration of pictorial semantics, from details to foundations— and not just in the theory of line reviewed here. Such developments will inevitably require a steep and steady descent from the relative simplicity of definitions like that of L into nuance and complexity.[95] On the other hand, mathematically precise accounts of ordinary drawing systems, like that of DeCarlo et al. (2003), show that pictorial semantics is not merely an analytic contrivance practiced on artificial examples. Despite the wide-spread skeptical attitude that depiction is basically unsystematic, such results show the potential for a progressive and rigorous analysis of natural systems of depiction. It remains to be seen whether this method can be extended to the myriad systems of depiction in use across history and culture. But the hypothesis of this paper is that it can.[96] I hope to have shown that this hypothesis is at least plausible.

4 Conclusion

Carnap (1963, 938) once described language as "an instrument... like any other... useful for a hundred different purposes." Arguably the most basic such purpose is the mobilization of public signs to disseminate private beliefs. According to the the Pictorial Semantics Hypothesis, defended in the first half of this paper, systems of depiction embody a parallel technology of communication. Both languages and systems of depiction are based on *semantics*: systematic and conventional mappings from signs to representational content. By contrast, the thesis of the second half of this paper, the Geometrical Projection Theory, suggests that these semantics are profoundly divergent. Whereas the semantics of language are based on arbitrary associations and rules of *logical* transformation (intersection, function application, and the like) the semantics of depiction, on this account, is based on non-arbitrary rules of *geometrical* transformation.

These conclusions provide compelling evidence that semantic analysis can thrive outside of the linguistic arena. They suggest that a general semantics, understood as the inclusive study of signs, is a viable endeavor. Such a study would apply the method of semantic analysis across the representational spectrum, covering languages, diagrams, pictures, and all the varying complexity in which these modes can be combined. It would bring into the light phenomena rarely studied,

[95] For example, Pirenne (1970, 116-35) and Hagen (1986, 170-1) document systems in which the operative method of projection varies according to the *type* of object being depicted.

[96] Thanks to Tendayi Achiume for helping me see both the limitations and import of the present work in this way.

and shed new light on those which have come to seem familiar.

References

Abell, C. 2007. Pictorial Realism. *Australasian Journal of Philosophy* 85 (1): 1–17.

———. 2009. Canny Resemblance. *The Philosophical Review* 118 (2): 183–223.

Abell, C., and K. Bantinaki. 2010. *Philosophical perspectives on depiction.* Oxford University Press.

Arnheim, R. 1974. Hansen's curvilinear perspective. *The Journal of Aesthetics and Art Criticism* 32 (3): 424–424.

Bach, K. 1970. Part of what a picture is. *The British Journal of Aesthetics* 10 (2): 119–137.

Balog, K. 2009. Jerry fodor on non-conceptual content. *Synthese* 170 (2): 311–320.

Beardsley, M.C. 1958. *Aesthetics: problems in the philosophy of criticism.* San Diego, CA: Harcourt, Brace / Company.

Block, N. 1982. *Imagery.* Cambridge, MA: MIT Press.

———. 1983. The photographic fallacy in the debate about mental imagery. *Nous* 17 (4): 651–661.

Blumson, B. 2007. Resemblance and representation. PhD diss., Australian National University, Canberra.

———. 2009a. Defining depiction. *British Journal of Aesthetics* 49 (2): 143–157.

———. 2009b. Pictures, perspective, and possibility. *Philosophical Studies* 149 (2): 135–151.

Borg, E. 2006. Intention-based semantics. *The Oxford Handbook of Philosophy of Language:*250.

Burge, T. 2010. *Origins of objectivity.* Oxford: Oxford University Press.

Carnap, R. 1963. *The philosophy of rudolf carnap.* Ed. P.A. Schilpp. Open Court Publishing Company.

Carroll, N. 1999. *Philosophy of art: a contemporary introduction.* Psychology Press.

Chomsky, N. 1965. *Aspects of the theory of syntax.* The MIT press.

Clowes, M.B. 1971. On seeing things. *Artificial Intelligence* 2 (1): 79–116.

Cohen, J., and C. Callender. 2006. There is no special problem about scientific representation. *Theoria* 21 (55): 67–85.

Cummins, R. 1996. *Representations, targets, and attitudes.* Cambridge, MA: MIT Press.

Da Vinci, L., J.P. Richter, P. Taylor, and RC Bell. 1970 [c. 1500]. *The notebooks of leonardo da vinci.* Mineola, NY: Dover Publications.

DeCarlo, D., and M. Stone. 2010. Visual explanations. In *Proceedings of the 8th international symposium on non-photorealistic animation and rendering,* 173–178. ACM.

DeCarlo, D., A. Finkelstein, S. Rusinkiewicz, and A. Santella. 2003. Suggestive contours for conveying shape. *ACM Transactions on Graphics (TOG)* 22 (3): 848–855.

Descartes, R. 2001 [1637]. *Discourse on method, optics, geometry, and meteorology.* Ed. P.J. Olscamp (trans.) Indianapolis: Hackett Publishing.

Dretske, F. 1983. Knowledge and the flow of information.

Dummett, M. 1996. What is a theory of meaning? (i). *The Seas of Language* 1 (9): 1–34.

Durand, F. 2002. An invitation to discuss computer depiction. In *Proceedings of the 2nd international symposium on non-photorealistic animation and rendering,* 111–124. ACM.

Eco, U. 1979. *A Theory of Semiotics.* Bloomington, IN: Indiana University Press.

Feldman, J. 1999. Does vision work? towards a semantics of perception. In *What is cognitive science?* Ed. E. Lepore and Z. Pylyshyn, 208–229. Basil Blackwell.

Field, H.H. 1978. Mental representation. *Erkenntnis* 13 (1): 9–61.

Files, C. 1996. Goodman's rejection of resemblance. *The British Journal of Aesthetics* 36 (4): 398–412.

Fodor, J. 1975. *The language of thought.* Cambridge, MA: Harvard University Press.

———. 1990a. A theory of content, i: the problem. *A Theory of Content and Other Essays:*51–87.

———. 1990b. A theory of content, ii: the theory. *A Theory of Content and other Essays:*89–136.

———. 2007. The revenge of the given. In *Contemporary debates in philosophy of mind,* ed. B.P. McLaughlin and J. Cohn. Blackwell.

Fodor, J., and E. Lepore. 2002. *The compositionality papers.* Oxford University Press, USA.

French, S. 2003. A model-theoretic account of representation (or, i don't know much about art – but i know it involves isomorphism). *Philosophy of Science* 70 (5): 1472–1483.

Gibson, J.J. 1954. A theory of pictorial perception. *Educational Technology Research and Development* 2 (1): 3–23.

———. 1960. Pictures, perspective, and perception. *Daedalus* 89 (1): 216–227.

———. 1971. The information available in pictures. *Leonardo* 4 (1): 27–35.

———. 1978. The ecological approach to the visual perception of pictures. *Leonardo* 11 (3): 227–235.

Giere, R. 2004. How models are used to represent reality. *Philosophy of Science* 71 (5): 742–752.

Gombrich, E.H. 1960. *Art and illusion: a study in the psychology of pictorial representation.* London: Phaidon.

Goodman, N. 1968. *Languages of art: an approach to a theory of symbols.* Indianapolis: Bobbs-Merrill.

Greenberg, G. 2010. Beyond resemblance. Unpublished manuscript. Available at: `http://eden.rutgers.edu/~gabrielg`

———. 2011. The semiotic spectrum. PhD diss., Rutgers University.

Grice, P. 1957. Meaning. *Philosophical Review* 66.

———. 1975. Logic and conversation. In *Syntax and semantics, volume 3: speech acts,* ed. Peter Cole and Jerry Morgan, 41–58. New York: Academic Press.

Groenendijk, J., and M. Stokhof. 1991. Dynamic predicate logic. *Linguistics and philosophy* 14 (1): 39–100.

———. 1997. Questions. In *Handbook of logic and language,* ed. J. van Bentham and A. ter Meulen. http://staff.science.uva.nl/ stokhof/q.pdf: Elsevier/MIT Press.

Gunderson, Keith, ed. 1975. *Language, mind and knowledge.* Vol. 7. Minnesota Studies in the Philosophy of Science. Minneapolis: University of Minnesota Press.

Gunther, Y.H. 2003. *Essays on nonconceptual content.* The MIT Press.

Hagen, M.A. 1986. *Varieties of realism: geometries of representational art.* Cambridge: Cambridge University Press.

Hamblin, C.L. 1958. Questions. *Australasian Journal of Philosophy* 36 (3): 159–168.

Hansen, R. 1973. This curving world: hyperbolic linear perspective. *The Journal of Aesthetics and Art Criticism* 32 (2): 147–161.

Hansen, R., and J.L. Ward. 1977. Perception of pictorial space in perspective pictures. *Leonardo* 10 (3): 261–262.

Haugeland, J. 1991. Representational genera. In *Philosophy and connectionist theory*, ed. W.M. Ramsey, S.P. Stich, and D.E. Rumelhart. Erlbaum.

Helmholtz, H. 1962 [1867]. *Treatise on physiological optics.* Ed. J.P.C. Southall. Vol. III. Mineola, NY: Dover.

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

Howell, R. 1974. The logical structure of pictorial representation. *Theoria* 40 (2): 76–109.

Huffman, D.A. 1971. Impossible objects as nonsense sentences. *Machine intelligence:*295.

Hyman, J. 1995. Perspective. In *A Companion to Aesthetics.* Blackwell Publishers.

———. 2006. *The objective eye: color, form, and reality in the theory of art.* Chicago: University of Chicago Press.

Jakobson, R. 1987 [1921]. On realism in art. In *Language in Literature.* Cambridge, MA: Harvard University Press.

Kaplan, D. 1968. Quantifying in. *Synthese* 19 (1-2): 178–214.

———. 1989. Demonstratives. In *Themes from kaplan,* ed. Joseph Almog, John Perry, and Howard Wettstein. New York: Oxford University Press.

———. 1999. The meaning of ouch and oops.

Kennedy, J.M. 1974. *A psychology of picture perception.* Jossey-Bass Publishers, London.

King, J.C. 2007. *The nature and structure of content.* Oxford University Press, USA.

King, J.C., and J. Stanley. 2005. Semantics, pragmatics, and the role of semantic content. *Semantics vs. pragmatics:*111–64.

Kivy, P. 1984. *Sound and semblance: reflections on musical representation.* Princeton University Press.

Knight, H. 1930. Aesthetic experience in pictorial art. *The Monist* 40 (1): 74–83.

Kripke, S. 1977. Speaker's reference and semantic reference. *Midwest studies in philosophy* 2 (1): 255–276.
Kulvicki, J. 2006. *On Images: Their Structure and Content.* Oxford: Clarendon.
Lascarides, A., and M. Stone. 2009a. A formal semantic analysis of gesture. *Journal of Semantics* 26 (4): 393.
———. 2009b. Discourse coherence and gesture interpretation. *Gesture* 9 (2): 147–180.
Lewis, D. 1969. *Convention: a philosophical study.* Oxford: Blackwell.
———. 1970. General semantics. *Synthese* 22:18–67.
———. 1975. Languages and language. In Gunderson 1975, 3–35.
———. 1979. Attitudes de dicto and de se. *The Philosophical Review* 88 (4): 513–543.
———. 1986. *On the plurality of worlds.* Oxford: Basil Blackwell.
Loewer, B. 1997. A guide to naturalizing semantics. *A Companion to the Philosophy of Language:*108–126.
Loewer, B., and J.W.J.R. Godbey. 1978. Representational symbol systems. *Semiotica* 23 (3-4): 333–342.
Lopes, D. 1992. Pictures, styles and purposes. *The British Journal of Aesthetics* 32 (4): 330.
———. 1995. Pictorial Realism. *Journal of Aesthetics and Art Criticism* 53 (3): 277–85.
———. 1996. *Understanding Pictures.* Oxford: Oxford University Press.
———. 2009. Drawing in a social science: lithic illustration. *Perspectives on Science* 17 (1): 5–25.
Lubbock, T. 2006. Philippe de champaigne: vanité (1660s). *The Independent.* `www.independent.co.uk/arts-entertainment/art/great-works/de-champaigne-philippe-vanit-1660s-798027.html`.
Mackworth, A.K., and R. Reiter. 1989. A logical framework for depiction and image interpretation. *Artificial Intelligence* 41:125–155.
Malinas, G. 1991. A Semantics for Pictures. *Canadian journal of philosophy* 21 (3): 275–298.
Manns, J.W. 1971. Representation, relativism and resemblance. *The British Journal of Aesthetics* 11 (3): 281–287.
McCloud, S. 1993. *Understanding comics: the invisible art.* Kitchen Sink Press.
Mi, X., D. DeCarlo, and M. Stone. 2009. Abstraction of 2d shapes in terms of parts. In *Proceedings of the 7th international symposium on non-photorealistic animation and rendering,* 15–24. ACM.
Montague, R. 1970a. English as a formal language. *Linguaggi nella societae nella tecnica:*189–224.
———. 1970b. Universal grammar. *Theoria* 36 (3): 373–398.
Morris, C.W. 1946. *Signs, language and behavior.* Upper Saddle River, NJ: Prentice Hall.
Neander, K. 1987. Pictorial representation: a matter of resemblance. *The British Journal of Aesthetics* 27 (3): 213–226.
Necker, L.A. 1832. Observations on some remarkable optical phænomena seen in switzerland; and on an optical phænomenon which occurs on viewing a figure of a crystal or geometrical solid. *Philosophical Magazine Series 3* 1 (5): 329–337.
Newall, M. 2003. A restriction for pictures and some consequences for a theory of depiction. *Journal of Aesthetics and Art Criticism* 61 (4): 381–394.
Novitz, D. 1975. Picturing. *Journal of Aesthetics and Art Criticism* 34 (2): 145–155.
Partee, B. 1980. Semantics– mathematics or psychology? In *Semantics from different points of view,* ed. Bäuerle, Egli, and von Stechow. Springer-Verlag, Berlin.
Peacocke, C. 1987. Depiction. *The Philosophical Review* 96 (3): 383–410.
Peirce, C.S. 1906. Prolegomena to an apology for pragmaticism. *The Monist* 16 (4): 492–546.
Pinker, S. 1995. *The language instinct.* HarperPerennial New York, NY.
Pirenne, M.H. 1970. *Optics, painting & photography.* Cambridge University Press.
Putnam, H. 1981. *Reason, truth, and history.* Cambridge: Cambridge University Press.
Putnam, Hilary. 1975. The meaning of meaning. In Gunderson 1975, 131–93.
Pylyshyn, Z. 1999. Is vision continuous with cognition? the case for cognitive impenetrability of visual perception. *Behavioral and Brain Sciences* 22 (3): 341–365.

———. 2006. *Seeing and visualizing: it's not what you think.* Cambridge, MA: MIT Press.
Richard, M. 1981. Temporalism and eternalism. *Philosophical Studies* 39 (1): 1–13.
Rogers, B., and C. Rogers. 2009. Visual globes, celestial spheres, and the perception of straight and parallel lines. *Perception* 38 (9): 1295–1312.
Ross, J. 1997. *The semantics of media.* Kluwer Academic Pub.
Schier, F. 1986. *Deeper into pictures: an essay on pictorial representation.* Cambridge University Press.
Sedgwick, H.A. 1980. The geometry of spatial layout in pictorial represenation. In *The perception of pictures,* 33–90. Vol. 1. London: Academic Press.
Siegel, S. 2006. Which properties are represented in perception? In *Perceptual experience,* ed. Tamar Szabó Gendler and John Hawthorne, 481–503. Oxford: Oxford University Press.
———. 2009. The contents of perception. *Stanford Encyclopedia of Philosophy.*
Soames, S. 1989. Semantics and semantic competence. *Philosophical Perspectives* 3:575–596.
Sorensen, R. 2002. The art of the impossible. In *Conceivability and possibility,* ed. T.S. Gendler and J. Hawthorne, 337–68. Oxford University Press.
Stalnaker, R. 1984. *Inquiry.* Cambridge, MA: MIT Press.
———. 1997. Reference and necessity. *A Companion to the Philosophy of Language:*534–554.
———. 1998. What might nonconceptual content be? *Philosophical Issues* 9:339–352.
Stanley, J. 2000. Context and logical form. *Linguistics and Philosophy* 23 (4): 391–434.
Stanley, J., and Z. Gendler-Szabo. 2000. On quantifier domain restriction. *Mind and Language* 15:219–61.
Starr, William B. 2010. Conditionals, meaning and mood. PhD diss., Rutgers University.
Szabo, Z.G. 2010. Compositionality. *Stanford Encyclopedia of Philosophy.*
Tarski, A. 1944. The semantic conception of truth: and the foundations of semantics. *Philosophy and phenomenological research* 4 (3): 341–376.
Toker, F. 2003. *Falling water rising: frank lloyd wright, ej kaufmann, and america's most extraordinary house.* Knopf.
Tversky, B. 2000. Some ways that maps and diagrams communicate. *Spatial Cognition II:*72–79.
Veltman, F. 1996. Defaults in update semantics. *Journal of philosophical logic* 25 (3): 221–261.
Walton, K.L. 1990. *Mimesis as Make-Believe: On the Foundations of the Representational Arts.* Cambridge, MA: Harvard University Press.
Waltz, D. 1975. Understanding line drawings of scenes with shadows. In *The psychology of computer vision.* Citeseer.
Willats, J. 1997. *Art and representation: new principles in the analysis of pictures.* Princeton: Princeton University Press.
Wittgenstein, L. 1921/1997. *Tractatus logico-philosophicus.* Ed. D. Pears (trans.) and B. McGuinness (trans.) London: Routledge.
Wollheim, R. 1987. *Painting as an art.* Thames / Hudson.
———. 1998. On pictorial representation. *Journal of Aesthetics and Art Criticism* 56 (3): 217–226.

Appendix: Formal semantics for *L*

This appendix partially formalizes the semantics for perspective line drawing presented in §3. I begin by modeling the key notions of picture, scene, projective index, and perspective projection. These components are then integrated into definitions of accuracy and content for the system of perspective line drawing, which are subsequently shown to be fully compositional. My aim in what follows is to formalize the relevant notions only as much as is required to show that a fully precise and algebraic analysis is possible, but not so much as to provide every detail of such an analysis.

As we embark, it is necessary to remember that formal models, in linguistics and here, are by definition abstract, mathematical representations of real phenomena. Consider two prominent examples: (1) Linguists typically model *sentences* as abstract logical structures which they distinguish from physically instantiated *utterances*. Here too I will model *pictures* as abstract geometrical objects, distinct from physically instantiated *inscriptions*.[97] (2) Mathematical linguists define languages in part by infinitely large sets of sentences— not merely those that have actually been uttered, but all the possible sentences of that language. In the same way, I will define systems of depiction in part by infinitely large sets of pictures. From the perspective of naturalistic metaphysics, such formally-motivated definitions may seem problematic. But I flag these questions only to set them aside: the definitions here are intended only to give my account precise articulation while abstracting away from irrelevant details.

The analysis employs a basic geometrical framework. I will be discussing Euclidean spaces of 2 and 3 dimensions, of finite and infinite extent, which I will refer to as 2-SPACES and 3-SPACES respectively. For every dimensionality, there are infinitely many distinct spaces of that dimension, individuated by the chromatic properties of and spatial relations among the points that make up each space. A REGION is a part of a space, such as a point, line, line segment, plane, convex solid, and so on. A COLORED REGION is a region where every point in that region is associated with exactly one color, but I also allow for colorless regions.[98] An OBJECT is defined as a connected colored region. Finally, an EMBEDDING is a function which maps a region in one space to an isomorphic copy of that region in another space. The only kinds of embeddings relevant here are those that preserve metrical relations between points and their chromatic properties. An embedding may occur between spaces of the same or different dimension. For example, one embedding may locate lines of a particular 2-space within a particular 3-space; another may locate lines of a particular 2-space within a distinct 2-space.

Definition of picture and scene

Let us define a PICTURE as a colored, rectangular segment of plane in a 2-space.[99] The decision to define pictures in this way is obviously an idealization, for it rules out pictures which are round, pictures which are not perfectly flat, and so on. Perhaps most artificially, I assume that the color of regions in a picture can be characterized independent of viewers or viewing conditions.[100] We shall model SCENES in a parallel fashion. At the outset we defined scenes as time- and location-centered worlds. We will now assume that every such scene determines a 3-space, which may be populated by objects, or not, in any way. We will refer interchangeably to scenes, considered as time- and location-centered worlds, and the spaces they determine.

Definition of projective index

Recall that in order to construct a perspective projection of some scene one must specify both the position of a viewpoint from which the projection lines are drawn, and the position and orientation of a picture plane onto which the image of the scene is projected. The specification of these parameters is the job of the PROJECTIVE INDEX. Different kinds of projective indices are required for different methods of projection. In the case of perspective projection, the projective index consists

[97] Willats (1997, 98-100) argues that we should distinguish pictures from inscriptions on the ground that not all marks on a page are semantically significant. For example, there may be stray marks or smudges on a piece of paper that are irrelevant to interpretation. One can imagine an entire science devoted to identifying the correct mappings from marks on paper to lines in the picture. But such details only distract from the project of semantics, which is to provide general rules of interpretation.

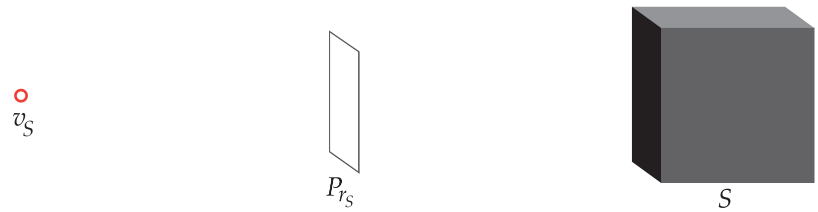
[98] Colors can be modeled as real numbers associated with positions on a color wheel.

[99] A richer representation of pictorial structure would parse pictures as arrangements of *parts*. Such an account would allow the theorist to classify "ambiguous" images like the Necker Cube as *syntactically* ambiguous, rather than merely semantically unspecific. (Necker 1832; Wittgenstein 1921/1997, §5.5423) Associating pictures with part structure may also play a role in intelligent interpretation of abstract images. (Mi, DeCarlo, and Stone 2009; DeCarlo and Stone 2010)

[100] This last assumption would be unsustainable in any careful study of color painting or photography. But here my goal is to take on as few complications as possible. Since I will only seriously examine systems of line drawing, such simplistic modeling choices will help, not hinder understanding. I hope that readers interested in color depiction can see how the definitions here could be extended to treat their preferred subject matter.

only of two components $\langle v, r \rangle$, where v is a VIEWPOINT FUNCTION, and r is a PICTURE POSITION FUNCTION.

The viewpoint function embeds a unique viewpoint— understood as a point— within the 3-space of a given scene. Formally, v is a function from scenes to points in the 3-spaces of those scenes. Where S is a scene, we shall refer to v applied to S, the viewpoint in S, as "v_S". The picture position function correspondingly embeds a picture— understood as a colored region within a 2-space— into the 3-space of a given scene. Formally, r is a function from pictures and scenes to colored flat regions in the 3-spaces of those scenes. Where S is a scene and P is a picture, we will refer to the picture embedded in S by r as "P_{r_S}". The following diagram shows a possible arrangement of the viewpoint and a picture P, as specified by v and r relative to the scene S.



Definition of perspective projection

I now define perspective projection in more exact terms. All forms of projection are defined in terms of an array of projection lines which pass through every point on the picture plane. Relative to this array of projection lines, methods of projection are defined by their PROJECTION and DENOTATION conditions. In §3.1, we described these conditions for the method of perspective projection as follows:

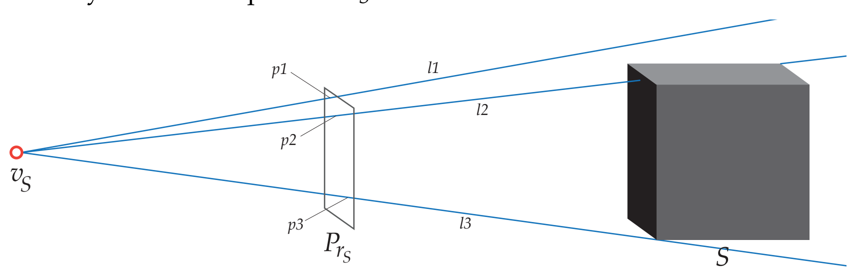
Projection: Every projection line intersects the viewpoint.
Denotation: A point in the picture is red if and only if it has a projective counterpart in the scene which lies on an accessible edge.

To formalize this proposal, we begin by defining the array of projection lines passing through every point on the picture plane, which these conditions presuppose. Where P_{r_S} is a picture situated relative to the space of the scene:

(23) for every point p in P_{r_S}, there is a projection line l intersecting p

We will treat these projection lines as rays. Then we can state the projection condition (i) as the requirement that all projection lines have their endpoints at the viewpoint v_S. That is, for any projection line l:

(24) l is a ray with its endpoint at v_S

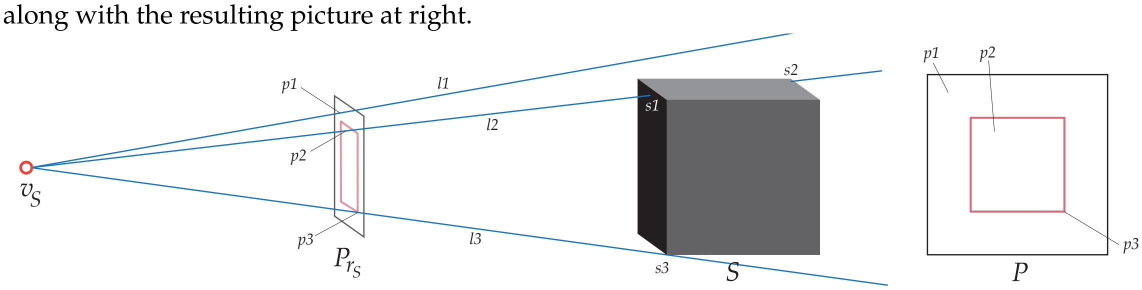


The diagram shows three projection lines $l1$, $l2$, and $l3$ passing through the picture plane at points $p1$, $p2$, and $p3$ respectively. Of course, these are only a representative selection of the total array; according to (23), some projection line passes through *every* point in the picture. As per (24), each projection line has its endpoint at the viewpoint. The interactions between these lines and the scene will become relevant shortly.

Next, the denotation condition (ii) specifies a rule for mapping edges in the scene onto lines in the picture plane. For line drawing (with red ink), the rule is that a point in the picture is red just in case it has some projective counterpart in the scene which is (a) accessible and (b) lies on an edge. Recall that a picture-point's "projective counterparts" are just those points in the scene which lie along the projection line which passes through that point. Then we may state the denotations condition as follows, where S is the scene, p is any point on the picture plane, and l is the projection line intersection p:

(25) p is red iff there is a point s on l such that:
(a) s is accessible from v_S along l;
(b) s lies on an edge in S.

A point in the scene s is ACCESSIBLE from the viewpoint v_S along a line of projection l just in case v_S and s can be connected by l without passing through any other point in the scene. An EDGE is the line segment at the intersection of two flat surfaces. (Note that I assume that every point in a picture is either red or white, thus the schema above determines the color for every point in the picture.[101]) The diagram below illustrates the application of (25) to each point on the picture plane, along with the resulting picture at right.



v_S P_{r_S}

Note how the three projection lines included here interact with the scene in different ways. First, $l1$ fails to intersect the scene at all; thus $p1$, the point at which $l1$ passes through the picture is left white. Next, $l2$ intersects the scene at two points; the first, $s1$ is accessible from the viewpoint, but does not lie on an edge; the second, $s2$, lies on an edge, but is not accessible from the viewpoint; thus $p2$, the point at which $l2$ passes through the picture is also left white. Finally, $l3$ intersects the scene at $s3$, which lies on an accessible edge; thus $p3$ is colored red.

Finally, by combining each of these components, we may define perspective projection. For any picture P, scene S, and projective index $j = \langle v, r \rangle$:

(26) The perspective projection of S relative to j, $\mathsf{proj}(S, j) = P$ iff
for every point p in P_{r_S}, there is a projection line l intersecting p such that:
(i) l is a ray with its endpoint at v_S;
(ii) p is red iff there is a point s on l such that:
(a) s is accessible from v_S along l;
(b) s lies on an edge in S.

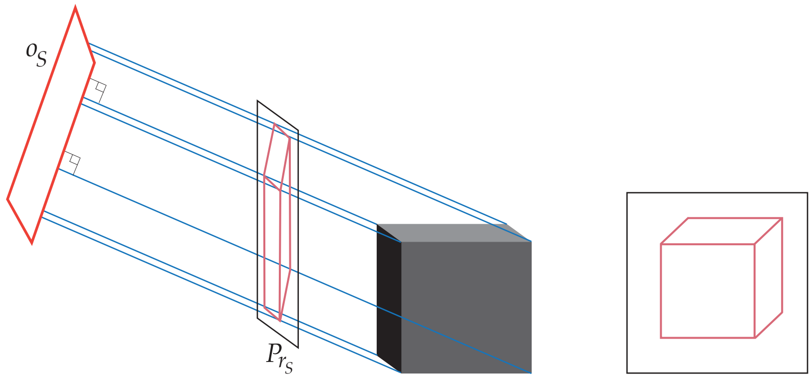
That is: a picture P is a perspective projection of a scene if and only if for each point p in P, as embedded in S by r, there is a projection line l such that (i) l originates at v_S and passes through p; and (2) p is red just in case there is some point s in the scene lying along l such that (a) s is accessible from v_S along l, and (b) s happens to lie on an edge in the scene.

By independently varying the projection and denotation conditions, one can derive a spectrum of projective methods besides that of perspective line drawing. Consider first the projection condition (i), which states that every point p in the picture defines a projection line l which intersects the viewpoint v_S and p. This simple stipulation characterizes all forms of perspective projection. Other

[101] The definition here fails if this assumption is relaxed, since on that condition, two pictures with different background colors would be compatible with the scene and index. Of course, the "whitespace" condition can be written into the definition of the projection function, but I find this distracting.

methods of projection define the projection lines in other ways. For example, in the method of parallel projection, the projection lines are defined as rays extended perpendicularly from a *plane*.[102] Letting o_S be such a plane, we could then write the projection condition for such as method as follows, where l is any projection line passing through a point in the picture P_{r_S}:

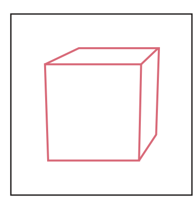
(27) l = the ray extending perpendicularly from o_S



The denotation condition (ii) describes how features of the scene determine features of the picture plane, given a certain projective mapping. The condition has two components. The first (a) states *which* of a picture-point's projective counterparts in the scene it is defined by; the second (b) states *how* it is defined by these counterparts. Alternatively, the first states *which* of a picture-point's projective counterparts it "says something" about; and the second states *what* it "says" about them.

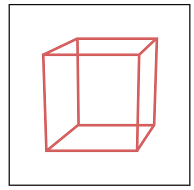
In most cases, (a) isolates those projective counterparts which are accessible from the viewpoint along the line of projection. Occasionally, however, this condition is modified to reveal occluded edges and faces. This results in the style of so-called "wire-frame" pictures like (29) below. To achieve this effect, we may weaken condition (ii) by eliminating the restriction to accessible points in S. The denotation definition then states that a point in the picture is red just in case it has *some* projective counterpart in the scene which lies on an edge. The second part of the denotation condition, (b), determines the relationship between the spatial or chromatic properties of a picture-point's projective counterparts, and the chromatic properties of the picture point itself. In the definition above, the picture-point is determined to be red just in case its accessible counterparts lie along an edge. But we may introduce some simple variations of this condition. To redefine the projection function as one that yields silhouettes, we may replace the denotation condition with (30). Then a picture-point is determined to be red just in case its accessible counterparts lie on *any* surface in the scene. To reveal the faces of the cube, but distinguish them from the edges, we must introduce two conditions— one for the edges, and one for the faces— as in (31).

(28) p is red iff there is a point s on l such that:
(a) s is accessible from v_S along l;
(b) s lies on an edge in S.

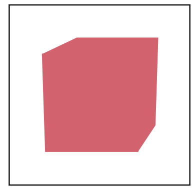


[102]Note that this would require changes to the definition of the projective index.

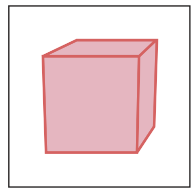
(29) p is red iff there is a point s on l such that:
(a) ——
(b) s lies on an edge in S.



(30) p is red iff there is a point s on l such that:
(a) s is accessible from v_S along l
(b) s lies on a surface in S.

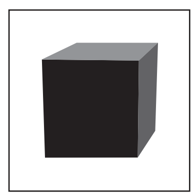


(31) p is red iff there is a point s on l such that:
(a) s is accessible from v_S along l;
(b) s lies on an edge in S.
p is pink iff there is a point s on l such that:
(a) s is accessible from v_S along l;
(b) s lies on a face in S.



These kinds of associations between surface type and point color are quite arbitrary. An edge could be mapped to a point of some other color, or something other than a point, such as a patterned region. But the relation between point color and properties of the scene may also be much more direct. For example, in a system reminiscent of color painting, the color of the points in the picture are simply inherited from the color of points in the scene:[103]

(32) for any color C:
p is C iff there is a point s on l such that:
(a) s is accessible from v_S along l;
(b) s lies on a surface in S and s is C.



In general, modifications to the denotation condition result in changes to the kind of information a picture records about its subject. Silhouette drawings are probably the least informative of the options just considered— they record merely the outline shape of a scene. Color projections contain the same outline information, but add to this the colors of the accessible surfaces in the scene. In the original system of edge-to-line projection, the picture reflected no color information about the scene, but it did contain additional spatial information not explicitly recorded in the color image— to wit, the location of edges in the scene.

Finally, it must noted that the original definition of perspective projection from §3.1 results in images composed of true *lines*, colored regions of zero width. Ideally, however, the definition should allow us to adjust line thickness as a parameter.[104] There are various ways to achieve this; the approach adopted here will require a slight change to the overall definition. The idea is that a point is colored red— not, as before, if it has a projective counterpart on an appropriate edge— but instead, if some *nearby* point has a projective counterpart on an appropriate edge. The result is that, for every one of the points that was assigned red by the original definition, there now will be a cloud of nearby points also assigned red. The metric of nearness defines the size of the "pen nib". To formalize this suggestion, let $n(\cdot)$ be a NIB FUNCTION which takes a picture P and point p in P and returns a circular region of P which includes p. Different nib functions may specify the size of this region differently. We then revise the definition of perspective projection as follows. For any

[103] Willats (1997, ch. 6) argues that many forms of color depiction map color features of *light* from the scene to points on the picture plane, rather than intrinsic color features of the scene itself (if such there be).
[104] See Durand (2002, §6.3) for discussion.

picture P, scene S, and projective index j:

(33) The perspective projection of S relative to j, $\mathsf{proj}(S, j) = P$ iff
for every point p in P_{r_S}, there is a projection line l intersecting some p' in $n(P_{r_S}, p)$ such that:
(i) l is a ray with its endpoint at v_S;
(ii) p is red iff there is a point s on l such that:
(a) s is accessible from v_S along l;
(b) s lies on an edge in S.

From perspective projection, we now turn to semantics.

Definition of system of depiction

A SYSTEM OF DEPICTION has three components. It specifies (i) the set of all possible pictures belonging to that system; (ii) the set of projective indices compatible with that system; and (iii) an accuracy function. The first component, the set of all possible pictures in the system, effectively fixes the system's syntax. In principle, the definition of this set might be quite complex, but at least for the systems included here, it is not. For our purposes, a picture belongs to the system of red on white line drawing just in case every region of the picture is a red region of a given minimum diameter or a white region.[105]

The second component of a system of depiction describes the range of projective indices compatible with that system. As we've seen, different systems of depiction require different kinds of projective indices. In some cases, systems put only technical constraints on indices, like substituting a plane of projection for the viewpoint characteristic of perspective projection. Other constraints are more "stylistic": some systems differ only in the orientation of the picture plane relative to whatever is considered the "dominant" plane of the scene.[106]

Finally, the accuracy function simply formalizes the relation of accurate depiction. It maps a picture P, scene S, and index j to 1 if and only if P perfectly accurately depicts S at j. (Thus the accuracy function is the correlate of the valuation function familiar from the semantics of propositional logic.) In this paper, we are concerned only with perfect accuracy and its failure, but in principle one might set out to give conditions for all degrees of accuracy between 0 and 1. The main task in analyzing a system of depiction is to define its accuracy function; here we shall do this implicitly, by specifying necessary and sufficient conditions on perfectly accurate depiction.

Semantics for *L*

In §3.2 we presented a schematic definition of accuracy for the system of simple line drawing in perspective, L:

(34) P accurately depicts$_L$ S at j iff $\mathsf{proj}(S, j)$ is in $\mathsf{flex}(P)$.

And the correlate definition of content:

[105] Researchers in computer science have sometimes claimed that the constraints on well-formedness for line drawings are much more substantive. (e.g. Huffman 1971, 313) They have deemed apparently uninterpretable drawings, like the Penrose triangle to be ill-formed, and have produced intricate formal theories of well-formedness. But I think this diagnosis of the situation is premature. In such cases it is possible to explain this uninterpretability semantically, by the fact that there are no objects that such a picture could depict. My methodological assumption is that it is always preferable to explain uninterpretability as semantic impossibility, rather than by fiat of syntactic rule. If I am right, the results of research on impossible drawings from computer science are significant, but they are not what their discoverers claim them to be. Rather, this field is the analogue of proof theory for pictures. In logic, proof-theory attempts to isolate the syntactic transformations which will produce or preserve semantic properties such as contradiction, theoremhood, and contingency. Similarly, research into impossible drawings attempts to isolate graphic transformations which will produce or preserve the semantic properties of impossibility and possibility. (As Wittgenstein (1921/1997, §2.225) observed, drawings cannot be *a priori* true; thus they cannot be theorems.) Further discussion of this question must be postponed.

[106] As Hagen (1986, 242-7) notes, in some systems there are clear conventions about the definition of the dominant face, dependent on object type. Modeling such a system would require a richer representation of scenes than I have elected for here.

(35) the content of P in L relative to j = the set of scenes S such that: $\mathsf{proj}(S, j)$ is in $\mathsf{flex}(P)$.

To obtain a explicit semantic analysis of L, we need only substitute into each of these formulae the definition of the projection function $\mathsf{proj}(\cdot)$ provided above, along with the definition of the flexibility function $\mathsf{flex}(\cdot)$ given in §3.2:[107]

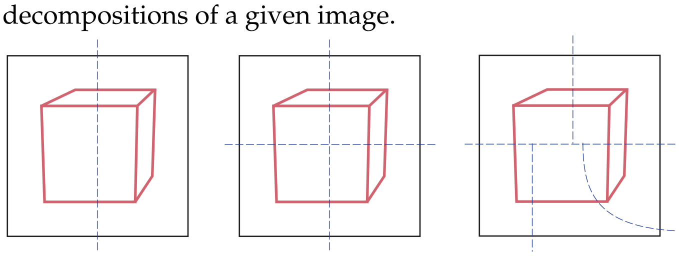
(36) $\mathsf{flex}(P)$ = the set of pictures which are scale duplicates of P.

Alternative systems of depiction can be derived by suitably redefining the projection function and flexibility functions. We have already considered several variations of the projection function corresponding to familiar alternative systems of depiction. Such variations can be multiplied by altering the flexibility function so as to allow some degree of freedom with respect, for example, to straightness of line or choice of color.

Compositional semantics for *L*

The semantics on offer here describes a general rule for interpreting pictures. The ability to interpret novel pictures is straightforwardly explained as the application of the general rule to novel particular cases. But it has been argued in the linguistic case that for this explanation to work, a semantics must be COMPOSITIONAL, that is, the content of the whole sign must be determined by the content of its parts and the way they are put together.[108] In this section I show how this principle may be realized for the case of pictures.

My treatment is shaped by two observations, both adapted from Fodor (2007, 108-9). First, I assume that every contentful part of a picture is also a picture. This is very different, note, from the linguistic case, where sentences have contentful non-sentential parts. A compositional semantics for pictures must show how the content of the whole picture can be derived from the content of pictures which are its parts. Second, I assume that, unlike complex linguistic expressions, pictures have no "canonical decomposition". That is, the content of a picture may be compositionally derived from *any* division of the picture into parts– so long as the parts are members of the same system as the whole. For any given picture, there are many possible decompositions; at the limit, at least for L, one may decompose a picture into points. Illustrated below are three arbitrarily selected possible decompositions of a given image.



To begin, let $D = \langle \mathbf{P}, \mathbf{m} \rangle$ be a DECOMPOSITION of a picture P relative to a system I. Intuitively, $\mathbf{P}$ is the set of picture-parts that results from cutting P up into pieces, and $\mathbf{m}$, which we'll call the ARRANGEMENT, shows you how to put them back together again. Formally, $\mathbf{P}$ is a sequence of finite, colored 2-spaces belonging to I; where the member of $\mathbf{P}$ at the nth position, is referred to as "p_n". Next, the arrangement $\mathbf{m}$ is a sequence of embeddings, each of which embeds a ranked member of $\mathbf{P}$ into the same 2-space; again the member of $\mathbf{m}$ at the nth position, is referred to as "m_n". $\mathbf{P}$ and $\mathbf{m}$ are coordinated in the following way: for every m_n in $\mathbf{m}$, $m_n(p_n)$ is a disjoint subset of P, and the union of every embedding in $\mathbf{m}$ applied to its corresponding member in $\mathbf{P}$ is identical to P— that is, $\bigcup_{i \le n} m_i(p_i) = P$.[109] The order of parts in $\mathbf{P}$ doesn't matter; but since it may be

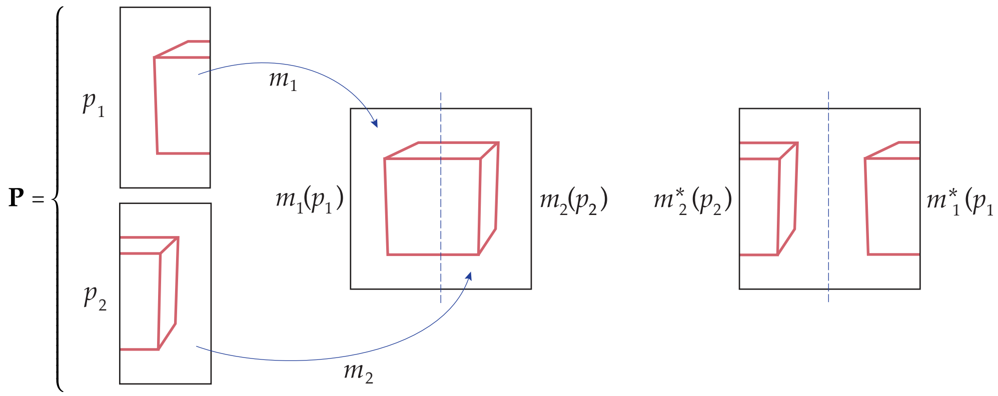
[107] The substitution is achieved perhaps more perspicuously with following equivalent reformulation of (36): P accurately depicts$_L$ S at j iff there is a P' in $\mathsf{flex}(P)$ such that $P' = \mathsf{proj}(S, j)$.

[108] See Fodor and Lepore (2002) and Szabo (2010).

[109] This way of defining the members of $\mathbf{m}$ requires that the spaces in question encode absolute position. To accommodate spaces that support only relative positions, $\mathbf{m}$ could be redefined so as to map *pairs* of members of $\mathbf{P}$ to relative locations in

necessary to map intrinsically identical regions to different locations, **m** distinguishes the different mappings by their order. According to the principle of compositionality, the content of the whole is determined by the content of the parts and the way they are combined; we will capture the way they are combined by the arrangement **m**.

The diagram below illustrates an example decomposition. Here the picture is divided into two parts such that $\mathbf{P} = \langle p_1, p_2 \rangle$ and the arrangement **m** maps these parts to their original positions within P. An alternative arrangement $\mathbf{m}^*$ maps the same parts to different positions, at right.



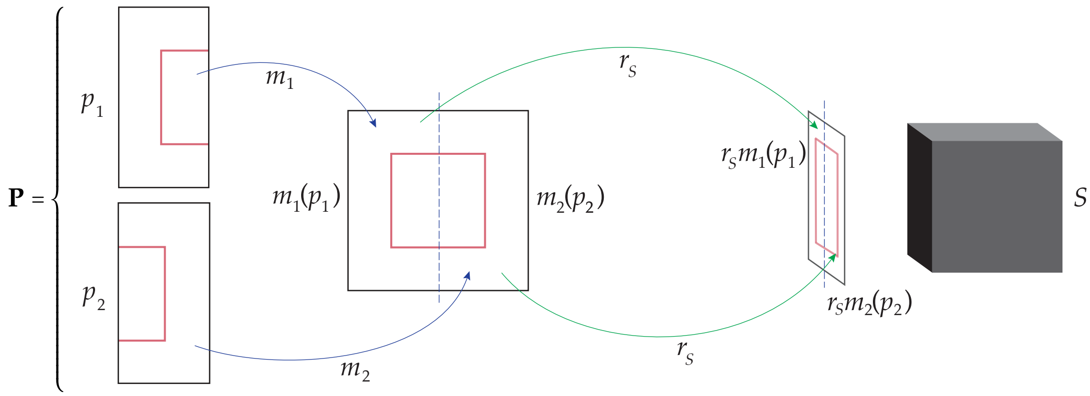
Having defined the relevant notions of a picture's parts and mode of composition, we now define the *content* of a part. Here we encounter a decision point: the content of a picture is defined relative to a projective index, which effectively fixes the vantage point of that picture. Is the content of picture's part defined relative to the same vantage point, or is it effectively "cut free" from the evaluative context of the original? Both seem to me reasonable ways of thinking about the content of a picture's parts. In what follows, I shall mainly elaborate the first approach, assuming that the projective index of the whole constrains the projective indices relative to which the the content of the parts are determined. Afterwards I will show how to execute a compositional analysis using the latter definition.

This tack engenders some technical complications. Recall that we defined the content of a picture relative to a projective index $j = \langle v, r \rangle$ where v is a viewpoint function and r is the picture position function. Given a scene S, r_S maps pictures in 2-space to flat regions in the 3-space of S. But we can't apply such a function straightforwardly to members of **P** in a given decomposition, for two reasons. First, there is a size mismatch: r_S takes whole pictures as inputs; the members of **P** are parts of pictures. The solution here is simply to let r_S range over regions of P, where the outputs are corresponding regions of $r_S(P)$.

A second challenge is more fundamental. The members of **P** are independent spaces, not regions of P; they contain no intrinsic information about their original relative position in P. Where in the scene then, should r_S embed such regions? How do we guarantee that they won't get misplaced in the process? The solution here is to first apply the embeddings in **m** to the members of **P**. Each such function cannily locates a member of **P** within a unified picture space. At that point r_S can be brought in to map subregions of this space to their corresponding regions in the space of S. The diagram below illustrates this technique for a simple example where $\mathbf{P} = \langle p_1, p_2 \rangle$.

---

a single 2-space.



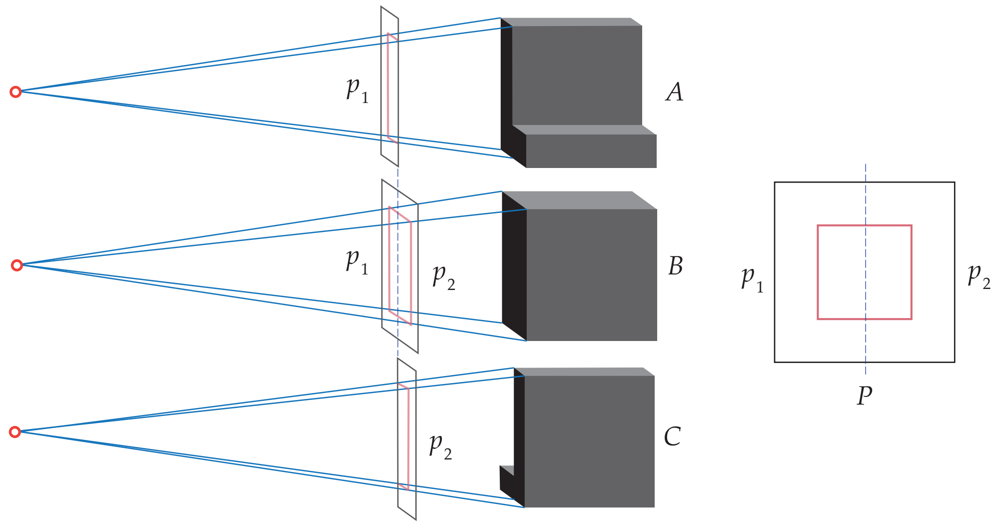
Given a member m_n of **m** and a picture position function r_S, we can define a function $r_S m_n$ which maps p_n into the space of S. (Correspondingly, rm_n takes a scene S and yields $r_S m_n$.) In this way we can say that the content of p_n is defined relative a projective index $\langle v, rm_n \rangle$. In general, where $j = \langle v, r \rangle$ is the projective index for a picture P, let $j_n = \langle v, rm_n \rangle$ be the projective index for any part p_n of P relative to a decomposition D. Thus we have arrived a definition of the content of a part of a picture.

In summary: we began with the idea of the content of a picture P in a system I at index j. We then introduced the concept of a decomposition D of P, which consists of an arbitrary division of the picture into parts, and a record of how these parts are combined in the original. Then, for every part p_n of P relative to the decomposition, we showed how to define the content p_n in I at index j_n, where j_n is determined by the decomposition D and the original index j. Thus given a picture, a system, an index, and a decomposition, we gave a general definition of the content of each part. To show that a pictorial semantics is compositional, we must further demonstrate that the content of the whole can be derived from the content of the parts and the way they are combined, as reflected in the decomposition. As we proceed, it will be convenient to refer to the content of a picture P in system L at an index j as "$[\![P]\!]_{L,j}$", and the content of a picture part p_n relative to decomposition D as "$[\![p_n]\!]_{L,j_n}$". What we aim to show is that, for any picture P in system L, index j, and decomposition D of P, we can derive $[\![P]\!]_{L,j}$ from $[\![p_n]\!]_{L,j_n}$ for every p_n in **P** and the mode of composition reflected in the arrangement **m**.

To motivate our final analysis, let us see how a simple first suggestion fails. According to this proposal, the content of a whole picture is the generalized *union* of the content of its parts. That is, given the system L, for any picture P in I and decomposition D:

(37) $[\![P]\!]_{L,j} = \bigcup\{[\![p_n]\!]_{L,j_n} \mid p_n \text{ in } \mathbf{P}\}$

Consider the predictions of this analysis for the case illustrated below. Here a picture P is decomposed into two parts p_1 and p_2. Image p_1 accurately depicts scene A and scene B, but not scene C. Image p_2 accurately depicts scene B and scene C, but not scene A. If we take the union of the sets of scenes that both p_1 and p_2 accurately depict, we get a set that includes all of A, B, and C. But obviously only B is in the content of P; P does not accurately depict A and C. Evidently, the content of P cannot be so defined; so (37) is false.



The failure of this method clearly suggests a better one. The content of P in the example above includes A, the scene accurately depicted by both p_1 and p_2, but none of the scenes accurately depicted by p_1 alone or p_2 alone. Thus it seems we should define the content of the whole, not as the generalized union of the content of its parts, but as their generalized *intersection*:

(38) $[\![P]\!]_{L,j} = \bigcap\{[\![p_n]\!]_{L,j_n} \mid p_n \text{ in } \mathbf{P}\}$

This analysis is evident enough, but here is an informal proof, with indices freely suppressed for brevity. Let D be a decomposition of a picture P into two parts, p_1 and p_2 (the number of parts is chosen arbitrarily for the example). Then proposition (38), applied to P is equivalent to the following claim: for any scene S, P accurately depicts S if and only if p_1 accurately depicts S and p_2 accurately depicts S. By the Geometrical Projection Theory, this in turn is equivalent to the proposition that P is a perspective projection of S if and only if p_1 is a perspective projection of S and p_2 is a perspective projection of S (here I suppress mention of the flexibility function). But this proposition follows directly. From left to right: if a picture can be derived from a scene by projection, then each of its parts can be so derived as well. From right to left: if all of a pictures parts can be individually derived from a scene when located at their original position in the picture, then the picture itself can also be derived from the scene. Thus (38) correctly defines the content of the picture in terms of the content of its parts.

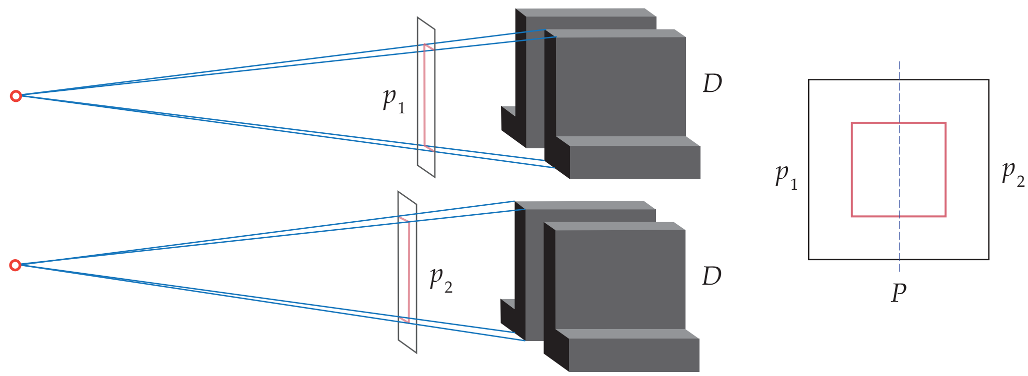
But is the proposed semantics *compositional*? Though it is clear how the content of the picture is derived from the content of its parts, one might wonder what happened to the "mode of composition" in this compositional semantics. In fact, it is encoded in the definition of the index j_n in terms of original index j and the arrangement $\mathbf{m}$. An unfortunate consequence of this approach that the conclusions derived above appear trivial; unlike the sophisticated combinatorics of language, the compositional rule for pictures amounts to little more than simple intersection. So it may prove more intellectually satisfying to see how compositionality can also be derived when the mode of composition is not built into the definition of part-content, but instead imposes a substantive constraint on how the content of the whole is derived from the content of its parts.

To this end, we will replace the definition of content at work in the preceding discussion, according to which the content of a picture is defined relative to a projective index, with the alternative conception of picture content discussed in footnote **??** which abstracts away from projective index, as follows:

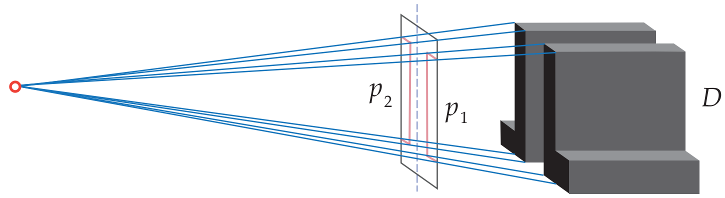
(39) $[\![P]\!]_L = \{\langle S, j\rangle \mid P \text{ accurately depicts}_L\ S \text{ at } j\}$

Here the content of a picture is modeled as the set of all scene-index pairs such that the picture accurately depicts the scene at that index. The definition applies to the content of parts of pictures, without modification. The result is that the content of a picture part is "cut free" from any projective index that might be used to evaluate the whole. As a consequence, picture parts do not encode their mode of composition in their content, and so there is no very simple technique for deriving the content of a whole picture from a set containing the content of each of its parts. As in the familiar linguistic case, merely intersecting the relevant contents is too unconstrained, leaving compositional semantics for pictures a substantive matter.

These observations are illustrated by the following case. In the diagram below, the scene D is incompatible with the content of the picture P, but compatible with each of its parts p_1 and p_2 considered individually— even when the viewpoint and orientation of the picture plane is held fixed across cases.



An unsophisticated method of intersection would incorrectly allow D (along with the common viewpoint) into the content of P, because it is in the content of both parts considered individually. What goes wrong with such a method? A clue is that the only conditions in which p_1 and p_2 can be projected from D at the same viewpoint are those in which they arranged relative to one another in a way that *differs* from their original arrangement in P. The diagram below makes this fact clear.



While the scenes in the content of the whole will indeed be contained in the intersection of the content of the parts, this case illustrates that such intersection is too permissive. Instead, we must specify a spatial arrangement of the contents so as to reflect the spatial arrangement of the parts in the original picture. This can be achieved by defining the content of the whole not directly in terms of the contents of its parts, but in terms of the content of its parts when their projective indices are suitably constrained to conform with their position in the picture's compositional structure. Thus for any system I, picture P and decomposition D of P:

(40) $[\![P]\!]_L = \{\langle S, \langle v, r\rangle\rangle \mid \forall p_n \text{ in } \mathbf{P} : \langle S, \langle v, rm_n\rangle\rangle \in [\![p_n]\!]_L\}$

That is: the content of a whole picture is the set of all scene-index pairs such that each pair is in the content of *every* part of the picture, when the index is suitably constrained to reflect the position of the part within the whole picture's compositional structure. Now more clearly than before, the content of the whole is determined by the contents of the parts ($[\![p_n]\!]_L$ for each p_n) and the way they are put together (by $\mathbf{m}$). Thus the proposed semantics for the system of perspective line drawing is fully compositional.