PHILOSOPHIA SCIENTIÆ

ROBERT SCHWARTZ

Pictures, puzzles, and paradigms

Philosophia Scientiæ, tome 2, n° 2 (1997), p. 231-242

<http://www.numdam.org/item?id=PHSC_1997__2_2_231_0>

Pictures, Puzzles, and Paradigms

Robert Schwartz
Department of Philosophy
University of Wiconsin Milwaukee

Philosophia Scientiae, 2 (2), 1997, 231-242

Abstract. In psychological theories of vision the projective paradigm dominates the study of picture perception. The symbolic paradigm, associated with the work of Nelson Goodman, offers an alternative account of pictorial representation and understanding. Adopting this latter perspective may help explain and resolve some of the puzzles plaguing research in the field.

Résumé. Dans les théories psychologiques de la vision, le paradigme projectif domine l'étude de la perception de l'image. Le paradigme symbolique, associé au travail de Nelson Goodman, offre un conception alternative de la représentation et de la compréhension picturale. L'adoption de cette perspective pourrait aider à expliquer et à résoudre certains des problèmes sur lesquels bute la recherche dans ce domaine.

Introduction*

When psychologists who study vision turn their attention to picture perception, they find themselves entangled in a web of puzzles. There is, moreover, no consensus and much confusion on how to resolve these matters experimentally. As a result, research on picture perception is in an uneasy state. When these same vision theorists turn their attention to Nelson Goodman's [1968] work on pictorial representation, they are highly critical. They are convinced his ideas are at odds with well-established facts. I think there is a connection between these two phenomena.

In brief, I believe Goodman and the vision theorists adopt strikingly different paradigms concerning the nature of pictorial understanding. Their disagreements, in the end, are less over the empirical data and more over the appropriate interpretation of the facts. At the same time, I believe the paradigm vision theorists do adopt is responsible for many of the puzzles they encounter. In what follows, I will use 'symbolic paradigm' to refer to the approach of Goodman and his followers, and 'projective paradigm' will serve to label the dominant paradigm of perceptual psychologists.

Grouping vision theorists in this way all under one rubric is, of course, a simplification. There are dissenters in the field who favor the symbolic model and other researchers who find neither model acceptable. In addition, there are significant differences among projectivists in the accounts of picture perception they champion. I think, however, these latter differences are mainly due to differences in their models of perception in general. The differences do not

---

* This paper is based on ideas further explored in "Two paradigms of picture perception: The uneasy state of research on picture perception", Report de Forschungsgruppe: *Perception and the role of internal regularities of the physical world* am Zentrum fuer interdisziplinaere Forschung der Universtaet Bielefeld, 1997.

indicate rejection of the projective paradigm's core conception of the nature of picture perception.

The Projective Paradigm

The basic idea of the projective paradigm is that seeing pictures involves the same psychological processes and mechanisms as seeing anything else in the world. In a sense this claim is trivial, since pictures are themselves physical objects in the world. The central projectivist claim goes further. Projectivists maintain that in an important psychological sense, seeing a representation of an object is like seeing the object itself.

Now in the case of seeing objects in the environment, the problem of perception may and is often conceived as being one of 'inverse optics'. Optics determines the projection of light rays from objects to the retina. In order to perceive the layout correctly, the perceiver must reverse the process. The perceiver somehow projects back from the retinal image, or the information contained therein, to the object from whence it came.

Vision theorists differ widely on how to explain this process. There is no agreement on the proper description of the stimuli, on the information available in the retinal image, on whether or what calculations are involved in recovering the scene from the image, and on much else. These are the sorts of differences, alluded to above, separating theorists who, nonetheless, adhere to the projective paradigm of picture perception. Where the paradigm's proponents agree is in assuming the propriety of adopting their favorite model of inverse optics to picture perception itself.

The guiding principle of the paradigm can be presented with the aid of 'Alberti's Window', a method for constructing 'realistic' pictures. As illustrated in numerous treatises on art and perception, the method requires placing a window between the artist and the scene to be depicted. The artist's task is to produce a picture that will duplicate the light rays at the point where they intersect the window on their way to the artist's eye. If a picture so constructed is then substituted for the window, it will project the same bundle of light rays to an observer's eye as the original object -- as long, that is, as the observer remains at the artist's original location, the so-called 'station point'. All this is simply a matter of optics.

According to the projective model, as the artist sees through Alberti's window to the object, so the viewer of pictures 'sees through' the picture surface and locates the represented scene in space. There is a continuity, so to speak, of the 'virtual' space depicted and the

environmental space perceived. 'Seeing through' is like 'seeing' the real scene except the source of the stimulus is not direct.

Implications

Once this projective paradigm is in place much else is taken to follow:

1. If perceiving pictures involves essentially the same processes and mechanisms as perceiving objects, then pictures can be used as substitutes for real objects in psychological experiments on vision. And such is common practice in visual research.

2. But, of course, in this context, the domain of countenanced pictures is highly restricted. It does not include many of the things we ordinarily call 'pictures'. No one thinks of using caricatures, ancient Egyptian, or Cubist pictures as substitute stimuli in experiments on, say, distance perception or shape perception.

3. More significantly for our concerns, the study of picture perception itself tends to be limited to this circumscribed domain. Only 'realistic' pictures, pictures constructed according to the rules of linear perspective, are assumed to fall within the scope of *visual* theory. Accounts of the understanding and cognitive role of other sorts of pictures, are considered tangential to perceptual theory. Why? Because it is hard to account for perceiving what they represent in terms of inverse optics.

4. As a first approximation, then, once the domain of pictures is so delimited, picture perception can be conceived along the lines of our everyday perception of the environment. In turn, the approach visual theorists take in explaining the perception of pictures depends mainly on the model of ordinary perception they adopt.

Puzzles

If, as projectivists assume, picture perception is of a piece with ordinary perception, how and why should there be any special puzzles about picture perception? Well, all theorists recognize one problem peculiar to pictures. Although most pictures represent three dimensional scenes, there is normally much information available indicating the picture itself is a flat surface. So it is claimed, a conflict exists in the visual stimuli pictures afford. There is a conflict between the two dimensional cues of the picture's own surface and the three dimensional pictorial cues. In some way the visual system must resolve such *cue conflicts* in order to perceive pictures. But how is this done?

On this matter there is little agreement. Various theorists propose models in which the perceiver suppresses or ignores the two

dimensional information. Others favor models which combine the two and three dimensional cues forming a compromise perception of the represented space. Another approach is to assume PURE picture perception is exhibited when or to the extent the two dimensional cues are eliminated or not available. As with the physicist's 'frictionless surfaces' or 'isolated systems', only in appropriately *idealized* set-ups is it possible to get at the real processes underlying the mechanisms at work. I think, the enormous experimental literature on picture perception involving monocular vision and other reduced viewing conditions, or in *trompe l'oeil* situations where the two dimensional cues are ineffective, attests to the influence of these ideas.

Of course, things get much worse once more realistic viewing conditions are considered. For it is not simply the presence of two dimensional cues that raises a problem. In most everyday situations, people are not located at the station point when viewing pictures. Unfortunately, inverse optics applied to the retinal images a picture makes available from these other viewpoints, does not project to the same scene or layout it does from the station point. Off the unique station point the stimulus array a picture affords is said to be *distorted.* This, though, raises deep questions about how perception can work when the stimuli are 'abnormal' and hence misleading.

Such distortions would pose less of a problem if perception were itself distorted in the way inverse optics predicts. And as Gombrich [1972] has pointed out, many theorists have adopted this 'curious myth'. A myth, Gombrich notes, because it flies in the face of ordinary experience. Pictures do not look terribly distorted when we move off the station point.

These days, few theorists maintain a very strong distortion thesis. It is generally admitted, for example, that a picture of the Cologne Cathedral is perceived, by and large, as representing the same view and shape of the building whether the picture is looked at from the station point or from a side. This fact, the resistance of perception to distortion, is attributed and referred to as the 'robustness' of perspective.

Robustness, while perhaps welcomed by the painter or photographer, is quite bothersome to the projectivist. For how can perception be robust when the stimuli are distorted? Examination of the picture perception literature would show this issue is a or *the* primary focus of current research. Here too, there is no agreement to its solution.

Some theorists deny the significance of robustness. They maintain picture perception is not robust if the observer is deprived of inappropriate information, in particular, cues indicating the

presence of the flat picture surface. PURE picture perception, again, is just inverse optics. Others hold the visual system takes into account the observer's location, recalibrates to the station point, and then solves the projection problem along usual lines. Gibsonians, eschewing 'taking account' models of perception in general, search for relevant higher-order stimuli, stimuli that remain invariant from one observation point to another.

Finally, and most distressing to numerous visual theorists, the distortion/robustness issue leads them to think it difficult, if not impossible, to make evolutionary sense of our ability to perceive pictures. Our visual system, after all, evolved to solve the projection problem in the everyday physical environment. Yet we readily perceive pictures under conditions in which the straightforward application of favored models of inverse optics break down. Since the ability to perceive pictures could not have had independent survival value, how, they wonder, could this capacity have ever evolved?

Thus once theorists adopt the projective paradigm puzzles abound. Among them are: a. cue conflict, b. cue distortion, c. robustness, and d. evolutionary coherence. There are, in fact, two other problems with the projective paradigm usually not recognized or ignored.

As mentioned, the inverse optics approach only seems plausible for a very small subset of what we ordinarily call 'pictures'. Caricatures, ancient Egyptian pictures, Cubists pictures and many more are not considered. Nevertheless, we readily perceive and understand these depictions. Their status remains most unclear. And the rationale for splitting them off and treating them separately from perspective pictures remains in need of adequate defense.

The projective paradigm provides, too, no ready means for dealing with various referential aspects of pictures. I have in mind here the sorts of issues Goodman presses at the beginning of *Languages of Art* in criticizing resemblance theories of representation. He shows the resemblance model cannot account for aspects of fictive representation, misrepresentation, or the mundane fact that identical twin brothers or the several prints of a lithograph are not ordinarily understood as representing one another. These features of pictorial representation do not seem to be explainable in projectivist terms.

Symbolic Paradigm

I assume everyone in this audience is familiar with Goodman's symbolic paradigm of representation, and I will not review it. I wish only to call attention to a few salient features of this approach. In

contrast to the projective model, the symbolic model assumes referential aspects of pictures are basic to their function. Thus, pictures are treated on analogy with languages as a form of symbolization. This idea was foreshadowed in Goodman's [1960] article "The Way the World Is". There he argued both that the picture theory of language is misguided and that adopting a language theory of pictures gives a better account of pictorial representation.

In *Languages of Art* Goodman extends the thesis. Pictures along with languages are just two of a very wide range of symbolic forms. Maps, gauges, music notation, graphs, diagrams, and the full range of what we ordinarily call pictures (caricatures, ancient Egyptian and Cubist pictures, etc.) are given a place.

Once the symbolic paradigm is in place much else follows. Switching the focus of the analysis in this way provides an alternative perspective on many of the puzzles plaguing the projectivist. It may, indeed, help resolve them. To begin, the symbolic paradigm provides a framework for handling issues of reference and misrepresentation, issues hard to handle while confined to the resources of the projective model. The symbolic paradigm, moreover, does not require the seemingly unmotivated constriction of the domain of pictures and pictorial perception. It offers, instead, a motivated basis for classifying symbolic systems, pictorial and non-pictorial, in terms of syntactic and semantic properties.

The symbolic paradigm also offers a different slant on the visual problems confronting and confounding the projectivist. Consider first the matter of cue conflict. The symbolic model sees no need to think of the cues caused by the flatness of the picture surface as in *conflict* with the three dimensional pictorial cues. The point is obvious in the context of other forms of symbolization. The sentence 'Cologne is on the Rhine' makes a claim about the environment, and in this sense has three dimensional significance. We do not, however, think the cues informing us of the sentence's status as a two dimensional written symbol in any way conflict with the three dimensional interpretation of its content. The symbolic paradigm suggests a similar account may be offered for perceiving pictures. We perceive a two dimensional pictorial symbol as having three dimensional significance.

Along similar lines, the symbolic approach may offer help with the distortion/robustness problem. Consider a sign bearing the sentence, 'The Cologne Cathedral is just ahead'. The sentence is about the Cathedral and offers information about its location. There is nothing perplexing, though, how this sign can be taken to represent these spatial relations when the sign is viewed from the side instead of straight-on. The stimuli and visual experiences of the written

sentence may change somewhat as we move about, but within limits we perceive the shapes of the letters correctly. Veridical perception of the written sentence, the representation, is all that is required to assess its content or meaning properly.

The symbolic paradigm suggests a similar approach to picture perception. A picture of the Cologne Cathedral may depict it as at a particular distance and having a particular size and shape. It makes no difference to this representational content whether the picture itself is viewed straight-on or from off its station point. True, the stimuli the picture affords change as we move about, and the perceptual experiences of the picture may differ to an extent. Yet, within limits, it is possible to perceive the shapes and relationships of the picture pretty much as they are. And that is what it takes to comprehend the picture's representational content.

The evolutionary dilemma projectivists confront is also given a new twist on the symbolic model. The locus of the problem is shifted, along with possible approaches to its solution. The paradigm suggests treating the issue not in isolation but in the context of other forms of symbolization. There is, for example, much controversy about the correct evolutionary account of the human language capacity. Yet no one supposes our ability to understand the meaning of written sentences is a deep problem for an evolutionary account of *vision*. Language comprehension depends on mastering the interpretive principles of the system. The failure of written words to replicate projectively what they represent does not stand in the way. Our ability to understand pictures may be best understood accordingly. Appreciation of the representational content of pictures requires having the requisite skills of interpretation. And disparities between the depiction and the depicted are no bar to this.

Humans do have an amazing, perhaps species defining, capacity to use many kinds of symbolic systems. Among the systems humans master are languages, graphs, and diagrams, systems whose representational schemes are relatively unconstrained. Other systems of representation, including mime, Greek sculpture, 'realistic' pictures, and for that matter ancient Egyptian pictures, are more systematic and in this way more constrained. Mastering the interpretative principles of these systems would appear the easier task. If this is so, their acquisition or development should pose less, not more, of an evolutionary quandary.

Reasons for Resistance

Given all the help the symbolic paradigm seems to offer the perceptual psychologist, why the reluctance to accept it?

I think this is primarily due both to a misreading of what the symbolic paradigm claims and to a prevalent assumption about the nature of vision. I will look at these each in turn.

Projectivists believe because the symbolic paradigm claims pictures function like languages, the model must and does claim pictures *are* languages. Projectivists, however, are convinced empirical evidence shows the mechanisms involved in 'reading' pictures, and the routes leading to the development of this skill, are not the same as those underlying the ability to read linguistic texts. Thus they find the symbolic paradigm untenable. (Such complaints are repeated over and over in criticism of *Languages of Art*). These complaints, though, rest on a misconception. The symbolist admits, indeed insists, depictional and linguistic systems differ in syntactic and semantic principles. Reading pictures, therefore, is not identical with reading words. But symbolists find here no basis for abandoning their paradigm. After all, as the above discussion makes clear, perceiving pictures typically is 'not exactly the same' as perceiving the real three dimensional environment. What's more, the simple dichotomy of symbol systems into pictures and languages is much too blunt. It leaves no obvious place for a range of other symbolic forms, maps, models, diagrams, music notation, and a whole lot more. The dichotomy serves to misdirect and obscure the study of the psychological mechanisms underpinning mastery and competence of these systems.

Projectivists tend to ignore such forms of representation and the issues they raise for a theory of perception. Instead, projectivists merely assume the major break among kinds of symbolic systems is between their chosen domain of realistic pictures and all the other types of description and depiction. This narrow class of depictions is thought to constitute a 'natural kind', *the* proper subject of investigation in the study of picture perception. But what is the rationale and motivation for this claim besides steadfast commitment to the paradigm? This leads to the second reason projectivists have for rejecting the symbolist's aid.

I think a formative intuition is the idea that understanding pictures is something our *visual* system does, without *cognitive* intrusion. Comprehension of other kinds of depictions and descriptions involve more than the visual faculty. Extracting the representational content of caricatures, or ancient Egyptian and Cubist pictures, like comprehending sentences in English, involves cognition. By contrast, it is not necessary to interpret realistic pictures. They are simply *seen*. Picture perception is something the visual system does without the intrusion of 'mental' interpretation.

The pervasiveness of this central intuition should not be

mistaken for clarity of formulation. There is no agreement among vision theorists as to what it means for a process to be mental and no consensus at all where vision leaves off and cognition begins. I have discussed this issue in detail elsewhere [1994] and can no more than allude to some of the problems most germane to our present concerns.

Often in discussions of the boundaries of vision, 'cognition' is equated with conscious deliberation, and picture perception is said to be free of such intrusion, hence, non-cognitive. This conception of cognition, however, can not serve to support the projectivist's intuition. For comprehending sentences is in this sense as non-deliberative or 'thoughtless' a process as understanding photographs. Yet language comprehension is supposed to be cognitive, going beyond what is *given* in perception.

Another prominent account of cognitive intrusion appeals to *learning*. In order to comprehend a sentence, we must learn the syntactic and semantic features of the language. Skill at extracting the representational content of real pictures is supposed to be different. It does not require experience or practice. The sway of this idea is reflected in the importance attached to claims that young children, or adults from distant cultures, comprehend perspective pictures without instruction.

This attempt to underwrite the core intuition also runs into difficulties. First, there is much dispute over the proper interpretation of the data on untutored picture perception. Second, evidence for untutored comprehension of perspective pictures must be understood in light of evidence showing comprehension of cartoons, caricatures, and other kinds of 'non-realistic' depiction may likewise not require explicit training. Third, in contemporary theories of vision the learned/innate distinction does not pair up with the cognitive/non-cognitive dichotomy supposedly underlying the core intuition [Schwartz 1994].

Finally, contrary to prevalent assumptions, I do not think the focus on learning truly gets at the heart of the projectivist's intuition. For suppose Latin were innate and required no learning to understand. The projectivist would still want to maintain Latin should be grouped with languages and not pictures. And the rationale would remain as before. Language comprehension is a two-stage process, seeing the words and then mentally interpreting them. Perceiving pictures is supposedly different. It is a one-stage process not requiring interpretation. We simply 'see through' pictures to the worlds they represent. There is no need for a second stage of interpretation.

Visual theory may explain seeing words, but surely it is no part of visual theory to account for how we determine what words

represent. In contrast, it is the job of vision, not mind, to perceive what pictures represent. Which pictures? Well only perspective pictures, the rest are to be lumped with languages.

The State of Research

The above account of the competing paradigms, I believe, sheds light on the uneasy state of research in picture perception. Usually in work on vision the symbolic framework is disregarded, for the problems it raises are thought to lie outside the scope of perception. If understanding a picture is like understanding a sentence, it is not a job for the visual scientist to investigate. At the same time, the highly circumscribed set of issues and domain the projectivist countenances make for a dubious research program. The projectivist studies only perspective pictures and only up to the point where vision ends and cognition begins. This puts the visual theorist in a bind.

If by severely restricting viewing conditions, the stimuli from picture and object can be made identical, as they are in various experimental set-ups, there is nothing really left to explain about *picture* perception. Once outside these non-standard confines, however, the stimuli afforded by pictures and their represented objects diverge, the more so as motion is allowed. Then there does seem to be distinctively pictorial phenomena for the visual scientist to investigate. But the greater the discrepancy between depiction and depicted, the less sense can be made of the projectivist's thesis. With each step beyond the limited domain of perspective pictures, the paradigm loses application. Thus the paradigm has nothing to say about the vast range of representations ordinarily classified as pictures.

A related tension lies in the formative intuition supporting the paradigm's delimitation of subject matter. The basis for claiming perspective pictures constitute a 'natural kind' for visual science gets its life from the assumption there is a significant demarcation between the products of vision and the products of mind. The comprehension of written language and non-realistic depictions is regarded as a two-stage process. Vision stops after generating an uninterpreted sentence or depictional display. Higher level cognitive mechanisms take over from there and extract the representational content. In the case of realistic pictures, the story is supposedly different. The representational content is extracted by the *visual* system. There is no need for a second stage.

Although this one-stage/two-stage distinction is easy to avow, it is not very easy to give it empirical content [Schwartz 1994]. In earlier times, matters were more straightforward. The sensory

domain was identified pretty closely with features thought to correspond to the retinal image. And not much processing was assumed to take place until central, cognitive centers of the brain were reached. Today we know there is selection, supplementation, and deletion beginning at the periphery and continuing to the end. The 'innocent eye' loses its innocence at the retina. So where is the projectivist to draw a well-motivated line?

On the one hand, the more *in*clusively the scope of the visual is conceived, the harder it is to exclude the perception of caricatures, Cubist pictures, and perhaps even sentences from its domain. This is not acceptable to the projectivist. On the other hand, a minimalist understanding of the visual raises opposite problems. A natural minimalist position might be to draw the boundary of the strictly visual at the extraction of 'basic' spatial information about the environment. This, however, threatens to collapse the projectivist's enterprise. To treat a flat painted surface as a *picture* requires more than seeing it as a colored object of a particular size, at a certain distance and direction. It must be perceived not simply as an object in the world but as a *representation.* Here commitment to the projective paradigm gets in the way. Inverse optics does not readily accommodate many of the important aspects of picture perception highlighted by the symbolic paradigm. And this I believe is a major reason for the uneasy state of research in picture perception. For stripped of 'interpretation', of 'reading', of the accretions of experience and all else that constitutes or contributes to referential and representational significance, a picture cannot function to guide behavior, inform cognition, or enhance aesthetic experience. Or in Goodman's terms, the projective paradigm has trouble accounting for the role pictures play in making and remaking our worlds.

References

Gombrich, E. H.

1972 The 'What' and the 'How': Perspective representation and the phenomenal world, *in* : R. Rudner & I. Scheffler (Eds.), *Logic and art*, Indianapolis : Bobbs-Merrill, 129-149.

Goodman, Nelson

1960 The way the world is, *Review of Metaphysics*, 14, 48-56.

1968 *Languages of art*, Indianapolis : Bobbs-Merrill.

Schwartz, Robert

1994 *Vision: Variations on some Berkeleian themes*, Oxford : Blackwell Publishers.