

## The Conflicted Character of Picture Perception

It is often said that picture perception has a dual character, which makes looking at pictures fundamentally different from looking at things like trees, chairs, and people. This duality can be explained in a variety of ways, but it is most commonly explained in terms of competing perceptions of space.<sup>1</sup> That is, it is frequently asserted that the perception of pictures (at least those depicting objects in space) involves two distinct impressions of space—the two-dimensional picture surface and the three-dimensional depicted scene—that are somehow in conflict, or inconsistent, with one another. For instance, R. L. Gregory contends that, visually, pictures have “an extraordinary double reality: flat objects seen as flat, and at the same time as quite different three-dimensional objects in a different space.”<sup>2</sup> Similarly, James J. Gibson claims that “a picture is both a surface in its own right and a display of information about something else. The viewer cannot help but see both, yet this is a paradox, for the two kinds of awareness are discrepant.”<sup>3</sup> In much of the psychological literature on the subject, this perceptual conflict is generally understood as a function of “conflicting cues.” When we look at a picture, so the story goes, there are cues that indicate that we are looking at a flat surface, and also cues that indicate that we are looking at a three-dimensional scene. These cues suggest two mutually incompatible impressions of the object we are looking at (the picture), which our visual system is able to resolve in one way or another. The end result is that we perceive both the two-dimensional surface of the picture and a three-dimensional “pictorial space” simultaneously, despite the inherent contradiction between the two.

In what follows, I will argue that the view that picture perception is dualistic, in the sense that it involves the reconciliation or accommodation of

two competing impressions of space, is mistaken. First, the notion that there is a contradiction in perceiving an object as flat while perceiving “pictorial space” ignores the fact that the two impressions of space at issue are fundamentally distinct. In order for there to be a contradiction, a picture would have to be flat and not-flat in the same respect, or in the same sense. However, the two-dimensional picture surface is perceived as belonging to “environmental” or “real space,” while the three-dimensional scene is perceived in terms of an “imaginary” or “pictorial space”—and therefore there is simply no contradiction in the simultaneous awareness of each. Second, the contention that picture perception involves confronting conflicting depth cues is based on an unwarranted and problematic assumption. The notion that we are presented with conflicting cues when we look at a picture presupposes that the same cues always communicate the same information, regardless of the context in which we encounter them. Yet, the nature of picture perception itself provides evidence that the information that a given cue communicates depends, at least to some extent, on the context, or its combination with other cues. Given these difficulties with the popular view that there is an inherent conflict in picture perception, then, we ought to reconsider the connection between perceiving a picture’s surface and perceiving “pictorial space.” Although there seems to be an important relationship between these two perceptual activities, this relationship should not be characterized as one of conflict.

### I. PICTURES AS PARADOXES

R. L. Gregory has provided perhaps the most extreme statement of the view that there is an

inherent conflict in picture perception. According to Gregory, “pictures are paradoxes” in that looking at a picture involves the simultaneous awareness of a surface and a three-dimensional scene: “no *object* can be in two places at the same time; no object can lie in both two- and three-dimensional space. Yet pictures are both visibly flat and three-dimensional . . . Pictures are impossible.”<sup>4</sup> For Gregory, the paradox that pictures pose to the visual system is not unlike other more obvious visual paradoxes: for instance, the inconsistent spatial features of M. C. Escher’s work, or the “impossible figures” that psychologists have created in order to conduct experiments. In fact, Gregory’s description of the effects of Schuster’s three-pronged ambiguous figure echoes his comments regarding the paradoxical nature of pictures in general: “Is there a middle prong? If so, does it follow the same position in depth as the outer two, or does it dip down below them? It seems, in a curious way, to do both at the same time. But that is impossible for one object, or one part of an object cannot exist in two places at the same time.”<sup>5</sup> All pictures of three-dimensional scenes, then, are allegedly paradoxical or impossible in the very same sense that Schuster’s three-pronged figure is paradoxical. In both cases, the visual system is presented with an object that possesses two mutually inconsistent spatial properties. Pictures of impossible figures are simply a special case of the paradox that all pictures represent—in such instances we are faced with an *additional* perceptual paradox. The ordinary pictures that we commonly encounter in our day-to-day lives may be more frequent and cause us less confusion, but they are no less paradoxical.

Gregory’s contention that pictures present the visual system with inconsistent impressions of space continues to have currency in current literature on picture perception. For instance, Reinhard Niederée and Dieter Heyer argue that the spatial duality of pictures represents a fundamental challenge to popular accounts of visual perception.<sup>6</sup> The problem, as they see it, is that these popular accounts require that for any given stimulus, the visual system must produce “a single consistent scene description.”<sup>7</sup> However, they argue that when we look at a picture, we perceive both a two-dimensional surface and a three-dimensional scene at the same time. Consequently, no single consistent scene description is possible, “for, how could a flat opaque surface on a wall and a

transparent opening in that wall with a scene being visible through it coexist at the same location?”<sup>8</sup> Similarly, Rainer Mausfeld suggests that the dual character of pictures consists in the fact that they “can generate an in-depth spatial impression of the scene depicted while at the same time appearing as flat two-dimensional surfaces hanging on a wall.”<sup>9</sup> Mausfeld, like Niederée and Heyer, accepts that there is a perceptual conflict here, in the sense that “we seem to have two mutually incompatible spatial representations at the same time.”<sup>10</sup> These authors are unlike Gregory in that they reject the notion that the perceptual conflict inherent in picture perception makes pictures somehow unique or peculiar, and they disagree as to how the visual system handles this perceptual conflict. However, there is fundamental agreement among these authors that perceiving a picture’s surface and perceiving pictorial space constitutes a contradiction, and that the fact that we are able to do both simultaneously requires some specific explanation.

This notion that looking at a picture involves mutually incompatible impressions of space may seem intuitively plausible. After all, it is quite true that “no object can lie in both two- and three-dimensional space,” as Gregory says. However, when one examines this claim closely, there appears to be an equivocation with respect to the meaning of “space.” In order for there to be a contradiction inherent in picture perception, we would have to see pictures as both two-dimensional and three-dimensional in the same sense. However, it seems clear that when we say that a picture looks two-dimensional, and also looks three-dimensional, we are talking about two very different kinds of space. The obvious difference between looking out a window and looking “into” a picture is that the space in the first case looks like the sort we could move around in, whereas the space in the second case does not. When we look at a picture we may be able to make judgments about whether certain actions could hypothetically be performed within the depicted scene; however, we obviously do not think we could perform any such actions within the pictorial space.<sup>11</sup> The fact that pictorial space is distinct from real or environmental space and is perceived as such is not something that those who find a contradiction in picture perception would deny. For example, Niederée and Heyer point out that the “planar aspect” of a picture “is experienced as belonging to ‘real space,’” while the “spatial aspect”

is not.<sup>12</sup> Similarly, Mausfeld comments that while a line drawing of a cube would be seen as a three-dimensional object, “it would hardly fool us into trying to grasp and rotate the cube.”<sup>13</sup>

Although these writers recognize the distinction between real and pictorial space, they do not recognize the full significance of the distinction; once it is granted that real and pictorial space are distinct, the supposed contradiction inherent in picture perception disappears. There is simply no contradiction or conflict in an object appearing to take up two-dimensions in real space while also appearing to take up three-dimensions in a pictorial or imaginary space. And, of course, there is not a single “object” being perceived as having two distinct spatial characteristics—strictly speaking, an artist does not “make the picture plane appear largely three-dimensional,” as Gregory suggests.<sup>14</sup> Even if we were to say that certain “segments” of a picture’s surface appear to be related to one another both two- and three-dimensionally, there is still no contradiction so long as the relevant segments are perceived as related two-dimensionally in real space, and three-dimensionally in pictorial space (which is obviously the case for most pictures, since if the segments making up the depicted objects were perceived to be related three-dimensionally in real space, we would experience a full-blown illusion of depth and might be tempted to grasp a drawing of a cube). The apparent contradiction between perceiving a picture’s surface and perceiving the three-dimensional scene it depicts, then, is not a contradiction at all because the surface and the depicted scene do not occupy the same sort of space. That is, the picture is both two-dimensional and not-two-dimensional, but not in the same sense. Thus, we have a ready answer to Niederée and Heyer’s question regarding how a flat opaque surface and transparent opening through which a three-dimensional scene is visible can coexist at the same location: the two are not perceived as occupying the same location. That the surface and the scene are not perceived as occupying the same space is precisely why looking at pictures generally does not produce the kind of confusion that we experience when looking at pictures of impossible figures (a point that Gregory largely ignores). When we look at Schuster’s three-pronged figure, for example, we are confused by the fact that the middle prong seems to exist simultaneously in two locations of *pictorial space* because such a state of affairs represents a

paradox. But the fact that we can see the depiction of the three-pronged figure as three-dimensional while simultaneously perceiving the picture’s surface does not confuse us because the figure is not perceived as having three-dimensions in real space.

Yet, it might still be objected that, in fact, we have direct experience of the incompatibility of our perception of the surface and the depiction. Mausfeld, for instance, suggests that part of the reason for claiming that we “have two mutually incompatible spatial representations” when we look at a picture is that these representations “are available internally and we can, without any effort, switch back and forth between them.”<sup>15</sup> This contention that the conflict between a picture’s surface and the scene it depicts is expressed through phenomenal “switches” between the perceptions of each is not unique to Mausfeld. In fact, E. H. Gombrich believed that the conflict between perceiving a picture’s surface and perceiving what it depicts is so great that he insisted it is impossible to perceive both simultaneously. Referring to a picture of a battle horse, Gombrich claims it is impossible “to ‘see’ both the plane surface and the battle horse at the same time . . . To understand the battle horse is for a moment to disregard the plane surface. We cannot have it both ways.”<sup>16</sup> Nor must one accept Gombrich’s rather extreme view in order to recognize a fundamental rivalry between our perception of a picture’s surface and what it depicts. For instance, Patrick Maynard suggests that we should not deny the importance of the 2D/3D rivalry present in much pictorial depiction, since observers are able to shift their attention between the features of a depicted scene and the features of the medium, and because artists commonly exploit 2D/3D perceptual conflicts.<sup>17</sup> However, while it is certainly true that when we look at a picture we can focus our attention almost exclusively on either the surface or the depicted scene, and that an artist has techniques by which to draw our attention to one or the other, these facts are simply not evidence of an inherent conflict between the perception of the two-dimensional surface and the three-dimensional scene. When we listen to an actor recite Shakespeare, we can focus our attention almost exclusively on either the rhythm or the meaning of Shakespeare’s dialogue. Moreover, the actor has techniques by which to draw our attention to either the rhythm or the meaning as he or she sees fit. Similarly, when we listen to

a jazz quartet we can focus our attention on any one member of the group as we please, excluding the others—and of course the band has techniques for drawing our attention to one member or another, during his or her solo. Clearly, there is no inherent contradiction or rivalry between perceiving the “rhythm” and “meaning” of dialogue, nor between hearing a saxophone and a piano; these activities are not “mutually incompatible,” but when we experience both simultaneously they can act as rivals for our attention. Consequently, we should not conclude from the fact that we can “switch” between perceiving a picture’s surface and perceiving the three-dimensional scene it depicts that somehow these perceptual activities are inherently incompatible.

## II. CONFLICTING CUES

Many researchers in the field of picture perception would not fully accept Gregory’s contention that pictures are paradoxes; however, much of the literature on the topic nonetheless assumes that seeing a three-dimensional scene “in” a two-dimensional surface involves a fundamental perceptual conflict. Most commonly, this conflict is explained in terms of conflicting depth cues. James E. Cutting and Manfredo Massironi, for example, contend that the spatial duality of pictures “is carried by ‘conflicting cues’ . . . which are not particularly common in the real world.”<sup>18</sup> This concept of conflicting cues assumes a specific picture of how we perceive spatial relations, where the visual system constructs its “percepts” based on a variety of different sources of information. For instance, Cutting and Peter M. Vishton identify nine sources of information regarding spatial layout: occlusion, relative size, relative density, height in the visual field, aerial perspective, motion perspective, convergence, accommodation, and binocular disparities.<sup>19</sup> These various cues provide different kinds of information: for example, occlusion provides only ordinal information, not information about the amount of depth, whereas binocular disparities may provide an absolute metric for distance.<sup>20</sup> Also, many of the cues vary with respect to their effectiveness at different distances: for example, occlusion has the same effectiveness at all distances, whereas the informativeness of binocular disparities decreases dramatically with distance.<sup>21</sup> According to this view, the perception of spatial

relations emerges from a complex and somewhat elastic process of combining the information communicated by different cues, depending on the weight of the cues at different distances.

The most important point to consider with respect to the subject at hand is that sometimes these cues can communicate contradictory information. Under controlled laboratory conditions, researchers often intentionally create such conflicts in order to observe which cues have the most influence over the visual system.<sup>22</sup> Such research has had a significant impact on the basic assumptions of those working in the field of picture perception. The popular account of picture perception is that when we look at a picture, “monocular” depth cues such as occlusion, relative size, relative density, height in the visual field, and aerial perspective tell us that we are looking at a three-dimensional object or scene. However, at the same time, the information we get from cues such as accommodation and convergence, binocular disparities, and motion perspective tell us that we are looking at a two-dimensional surface. Although there are disagreements with regard to how the visual system handles this contradictory information, perhaps the most common view is that “human vision involuntarily strikes some kind of compromise between the flatness of the picture surface and the relief due to monocular cues.”<sup>23</sup> Such a claim is the common explanation for the often repeated observation that pictorial depth is enhanced when a picture is viewed with a single eye from a fixed viewpoint, and compressed when viewed normally.<sup>24</sup> Conversely, Niederée and Heyer contend that when faced with the conflicting cues a picture presents, our visual system creates both a two-dimensional surface “subpercept” and a three-dimensional scene “subpercept,” which coexist and are somehow bound into a unitary experience.<sup>25</sup> Despite such disagreements, however, there is fundamental agreement regarding the fact that looking at a picture involves the perception of conflicting cues that the visual system must accommodate in one way or another.

Just like the assumption that pictures are perceived as possessing mutually incompatible spatial properties, the notion that pictures present the visual system with conflicting cues is intuitively appealing. However, there is an extremely problematic assumption at work here: we can maintain that pictures present us with conflicting depth cues only so long as we assume that the same cues always

communicate the same information, regardless of the context in which they are perceived. To hold such a view, we would essentially have to believe that our visual system attaches a single meaning to every visual cue and never reinterprets the significance of any such cue under any circumstances. However, a position that denied the visual system the ability to adapt to different situations or contexts would be extremely odd, to say the least. It is common sense that we interpret the same visual stimuli differently, depending on our beliefs about our surroundings, and there is simply no reason to deny that the same is true of our perception of pictures. To the contrary, knowing that something is a picture—an artifact made with a certain intention—changes the way it looks to us and, in fact, changes the way we look at it.<sup>26</sup> When we open a magazine or glance at a billboard, we are expecting to see pictures, and it is only reasonable to assume that such expectations influence our interpretation of the visual cues we find in such situations.

Most significantly, if we assume that the same cues are always understood by the visual system as conveying the same information, we will not be able to account for the perception of pictorial space at all. To say, for example, that when we look at a picture the visual system reaches a compromise between the monocular cues suggesting a three-dimensional scene and the stereo information suggesting a flat surface, does not explain why we perceive the three-dimensional scene in pictorial as opposed to environmental space. A strict compromise in such a case would be to split the difference between the two sets of cues and perceive the depicted scene as occupying a highly compressed real space—a space in which a sufficiently small person could move around. The very fact that we perceive pictorial space, then, demonstrates that the same cues *do not* communicate the same information in every circumstance. We simply would not perceive pictorial space if the monocular cues in a picture did not communicate different information to the visual system when perceived in combination with cues indicating the flatness of the picture surface.

Hence, the most reasonable explanation of picture perception is that when binocular disparities, accommodation, and convergence tell us that we are looking at a flat surface, monocular cues such as occlusion and relative size provide information regarding the spatial relations of depicted objects

in pictorial space. If this is correct, it means that whether these monocular cues provide information about depicted objects in pictorial space, or physical objects in real space, depends on the information provided by other visual cues present in a given circumstance.<sup>27</sup> That is, we ought to conclude from the nature of picture perception that certain depth cues provide different information to the visual system, depending on the context in which they are perceived. According to this account, then, picture perception does not involve a conflict between depth cues at all. When we look at a picture there are not two sets of cues providing competing information about a single object; rather, there is a certain combination of depth cues that makes it possible for us to perceive depicted objects related three-dimensionally in pictorial space. Moreover, not only do these considerations suggest that picture perception is not inherently conflicted, they suggest that seeing depth in pictures is not even a particularly special perceptual act. When we look at a picture, a certain combination of cues produces a certain effect—namely, the impression of a flat surface in real space depicting a three-dimensional scene in pictorial space—which is a process not fundamentally different from looking at things like trees, chairs, and people. In other words, the perception of pictorial space is accomplished in much the same way that the perception of real space is accomplished.<sup>28</sup>

Given that the monocular depth cues provide information about objects in pictorial space, it is ultimately rather difficult to see where a conflict between cues could possibly be located. In fact, this problem with the very concept of “conflicting cues” has not gone unnoticed by researchers. For example, Jan J. Koenderink and Andrea J. van Doorn voice reservations about using the term: “because the observers are always simultaneously aware of both the optical space that contains the photograph as a physical object and the pictorial space that exists as a completely disparate entity, cue conflict is perhaps an unfortunate term, as it suggests that different cues are in conflict with respect to the resolution of a single entity.”<sup>29</sup> Yet, Koenderink and van Doorn seem compelled to use the language of “cue conflict” in order to account for the fact that pictorial depth is enhanced when those cues identifying the picture surface are eliminated. This phenomenon of “monocular depth enhancement” is, of course, important

and should not be ignored. The existence of this phenomenon means that we should not claim that our visual system simply ignores the surface of a picture when we make determinations about spatial relations between depicted objects, as Cutting seems to suggest.<sup>30</sup> However, neither can we say that monocular viewing increases the perception of pictorial depth by eliminating conflicting information: if we accept that there is a conflict between cues in picture perception, we might be able to explain why space is compressed when we look at a picture under ordinary conditions, but we will thereby undermine our ability to explain how the perception of pictorial space is possible in the first place. Thus, 'cue conflict' is not just an unfortunate term; in the context of picture perception it is a completely uninformative and misleading term. Given the empirical evidence, the most that can be said is that our perception of a picture's surface has some influence on our perception of the three-dimensional scene or object that the picture depicts. In other words, all we know is that different combinations of depth cues produce different effects. However, there is no reason to believe that the perception of a picture's surface influences the perception of a depicted scene by presenting the visual system with contradictory information—in fact, there are good reasons to believe that there is no inherent conflict in picture perception at all.

BOYD MILLAR  
University of Toronto  
Department of Philosophy  
Toronto ON M5S 1A2  
Canada

INTERNET: boyd.millar@utoronto.ca

1. See, for example, Patrick Maynard, "Seeing Double," *The Journal of Aesthetics and Art Criticism* 52 (1994): 155–167.
2. R. L. Gregory, *The Intelligent Eye* (New York: McGraw-Hill, 1970), p. 51.
3. James J. Gibson, *The Ecological Approach to Visual Perception* (Hillsdale, NJ: Lawrence Erlbaum Associates, 1986), p. 282.
4. Gregory, *The Intelligent Eye*, p. 32.
5. Gregory, *The Intelligent Eye*, p. 57.
6. Reinhard Niederée and Dieter Heyer, "The Dual Nature of Picture Perception: A Challenge to Current General Accounts of Visual Perception," in *Looking Into Pictures: An Interdisciplinary Approach to Pictorial Space*, ed. Heiko Hecht, Robert Schwartz, and Margaret Atherton (MIT Press, 2003), pp. 77–98.

7. Niederée and Heyer, "The Dual Nature of Picture Perception," p. 80.
8. Niederée and Heyer, "The Dual Nature of Picture Perception," pp. 80–81.
9. Rainer Mausfeld, "Conjoint Representations and the Mental Capacity for Multiple Simultaneous Perspectives," in *Looking Into Pictures*, p. 20.
10. Mausfeld, "Conjoint Representations," p. 25.
11. On this point, see Sheena Rogers, "Perceiving Pictorial Space," in *Perception of Space and Motion*, ed. William Epstein and Sheena Rogers (San Diego: Academic Press, 1995), p. 122.
12. Niederée and Heyer, "The Dual Nature of Picture Perception," p. 82.
13. Mausfeld, "Conjoint Representations," p. 28.
14. Gregory, *The Intelligent Eye*, p. 52.
15. Mausfeld, "Conjoint Representations," p. 25.
16. E. H. Gombrich, *Art and Illusion: A Study in the Psychology of Pictorial Representation* (New York: Pantheon Books, 1960), p. 279.
17. Patrick Maynard, *Drawing Distinctions: The Varieties of Graphic Expression* (Cornell University Press, 2005), p. 214.
18. James E. Cutting and Manfredo Massironi, "Pictures and Their Special Status in Perceptual and Cognitive Inquiry," in *Perception and Cognition at Century's End*, ed. Julian Hochberg (San Diego: Academic Press, 1998), p. 139.
19. James E. Cutting and Peter M. Vishton, "Perceiving Layout and Knowing Distances: The Integration, Relative Potency, and Contextual Use of Different Information about Depth," in *Perception of Space and Motion*, ed. William Epstein and Sheena Roger (San Diego: Academic Press, 1995), pp. 69–117. Cutting and Vishton also make it clear that these cues provide the information they do only on the basis of certain assumptions (see pp. 81–94).
20. Cutting and Vishton, "Perceiving Layout and Knowing Distances," p. 93.
21. Cutting and Vishton, "Perceiving Layout and Knowing Distances," p. 99.
22. For an overview, see Mausfeld, "Conjoint Representations," pp. 23–24.
23. Jan J. Koenderink, Andrea J. van Doorn, and Astrid M. L. Kappers, "On So-Called Paradoxical Monocular Stereoscopia," *Perception* 23 (1994): 583–594; quote is from p. 593.
24. See, for example, Rogers, "Perceiving Pictorial Space," pp. 128–129; Jan J. Koenderink and Andrea J. van Doorn, "Pictorial Space," in *Looking Into Pictures*, pp. 265–266.
25. Niederée and Heyer, "The Dual Nature of Picture Perception," pp. 90–91.
26. See Maynard, *Drawing Distinctions*, pp. 109–111.
27. This is all that the case of picture perception establishes. It may be that this variability in the way cues are interpreted by the visual system is much more widespread, and it might even be possible that the type of information provided by a given visual cue could change radically depending on the circumstances, but picture perception itself provides no evidence of this.
28. Accordingly, the present proposal needs to be distinguished from Richard Wollheim's view that, while the perception of pictorial space is not characterized by conflict, it nonetheless involves a special sort of perceptual

capacity. Wollheim claims that perceiving pictorial space involves a special ability that he calls “seeing-in,” which is exercised in our seeing human figures in stained walls or clouds, but that is not involved in our perception of ordinary objects and their spatial relations; see his *Art and Its Objects*, 2nd ed. (Cambridge University Press, 1980), pp. 217–219 and *Painting as an Art* (Princeton University

Press, 1987), pp. 46–47. Conversely, the present suggestion is that the perception of depth in pictures is merely an extension of the capacity to perceive depth in ordinary circumstances.

29. Koenderink and van Doorn, “Pictorial Space,” p. 265.

30. James E. Cutting, “Reconceiving Perceptual Space,” in *Looking Into Pictures*, p. 225.