Psychological Review
2002, Vol. 109, No. 3, 492–519

0033-295X/02/$5.00 DOI: 10.1037//0033-295X.109.3.492

Toward a Perceptual Theory of Transparency

Manish Singh and Barton L. Anderson
Massachusetts Institute of Technology

Theories of perceptual transparency have typically been developed within the context of a physical model that generates the percept of transparency (F. Metelli's *episcotister* model, 1974b). Here 2 fundamental questions are investigated: (a) When does the visual system initiate the percept of one surface seen through another? (b) How does it assign surface properties to a transparent layer? Results reveal systematic deviations from the predictions of Metelli's model, both for initiating image decomposition into multiple surfaces and for assigning surface attributes. Specifically, results demonstrate that the visual system uses Michelson contrast as a critical image variable to initiate percepts of transparency and to assign transmittance to transparent surfaces. Findings are discussed in relation to previous theories of transparency, lightness, brightness, and contrast–contrast.

Our visual world consists of coherent objects and surfaces distributed in a three-dimensional environment. The inputs to our visual systems, however, contain no such objects or surfaces—only two-dimensional arrays of light intensities. These light intensities result from a combined effect of many variables, including surface reflectance, the intensity and position of light sources, relative surface orientation, and characteristics of the optical medium. As a result, identical surfaces can project very different images onto our retinas under different viewing conditions. In order to compute the intrinsic attributes of objects and surfaces, the visual system must decompose the pattern of light intensities into the separate contributions responsible for the image data. This task is made enormously difficult by the fact that any given pattern of image intensities could have been generated by infinitely many distinct combinations of these various sources. A fundamental goal of vision science is to understand how the visual system performs this decomposition and assigns intrinsic properties to surfaces.

One of the most compelling forms of image decomposition arises in the perception of transparency. When observers view an object or surface through a partially transmissive surface (such as a mesh or filter), they are visually aware of two separate surfaces along the same line of sight. Although the information from these two surfaces has been collapsed onto a single intensity value at each location on the retina, the visual system is able to decompose the image and extract the two surfaces. The perceptual vividness of this decomposition suggests that transparency is likely to serve as an excellent subdomain to study the more general problem of image decomposition. Indeed, the perception of transparency has intrigued visual researchers since Helmholtz (1866/1962), who spoke of seeing one color through another, and Koffka (1935), who referred to the problem as one of *scission,* or splitting image intensities into multiple contributions. Figure 1, b and c, demonstrates the problem of scission[1] in perceptual transparency. In Figure 1b, the light gray patch in the central region is seen as a combination of two different colors: a transparent midgray over an underlying white surface. Similarly, the dark gray patch is seen as a combination of a transparent midgray and an underlying black surface. If the same patches are separated in the image, as in Figure 1c, the perception of transparency is lost, and each patch is seen as a single opaque color. What are the geometric (or figural) and photometric (or luminance) conditions that must be present to initiate the decomposition of image luminance into two surfaces layered in depth? How does the visual system assign surface properties (such as transmittance and lightness) to a transparent layer when such decomposition occurs?

Historically, answers to these questions have been dominated by Metelli's (1970, 1974a, 1974b, 1985) *episcotister* model of transparency. As we discuss in detail below, this model was developed within a specific physical context that elicits a percept of transparency. Metelli used the model to motivate a perceptual theory of transparency that identified perceptual scission with the inverse of the equations derived in this physical model. The influence of this model has been so compelling that the conditions for transparency are often referred to as the *Metelli conditions.* In part, this is attributable to the fact that most empirical work on perceptual transparency (even up to the present; see, e.g., Kasrai & Kingdom, 2001) has provided support for the validity of Metelli's model as

Manish Singh and Barton L. Anderson, Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology.

Earlier versions of the experiments reported in this article were presented at the annual meetings of the Association for Research in Vision and Ophthalmology, Ft. Lauderdale, Florida, May 1999 and May 2000, and the annual meeting of the Psychonomic Society, Los Angeles, November 1999.

Correspondence concerning this article should be addressed to Manish Singh, who is now at the Department of Psychology, Rutgers University—New Brunswick Campus, 152 Frelinghuysen Road, Piscataway, New Jersey 08854-8020. E-mail: manish@ruccs.rutgers.edu

[1] Consistent with Koffka's (1935) usage, we use the word *scission* to encompass not only instances of transparency but, more broadly, all cases in which two sources must be inferred from a single image intensity. Transparency is a special case of scission—albeit a unique one in that it involves inferring two distinct surfaces or material layers, rather than, say, a surface and an illuminant. Moreover, in this article, we use *scission* as a purely functional term, rather than as connoting any specific mechanism or as an explicit description of perceptual phenomenology.

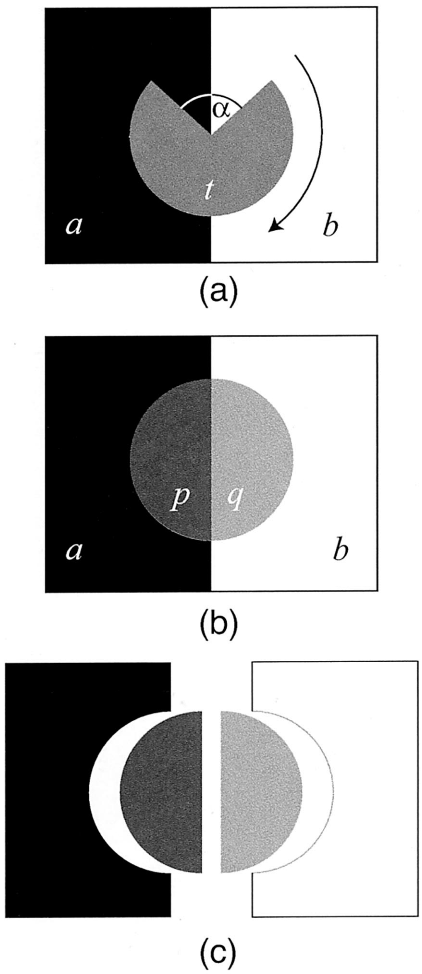

Figure 1. Metelli's episcotister model of transparency. (a) A disk with reflectance t and an open sector of relative area α is rotated rapidly in front of a bipartite background with reflectances a and b. (b) When this rotation is sufficiently fast, it leads to the percept of a transparent surface overlying the bipartite background. This generates reflectances p and q in the region of perceived overlay. (c) If the four regions are separated from each other, a perceptual decomposition into two separate layers does not occur, and each of the two gray patches is seen as a single opaque color.

a theory of perception (but see Beck, Prazdny, & Ivry, 1984). Ironically, however, Metelli himself had realized that something was wrong with the perceptual predictions of his model. Metelli (1974a) noted that a black episcotister looks more transmissive than a white episcotister of the same physical transmittance,[2] whereas his model predicts that they should look equally transmissive. Thus, despite its apparent successes, this very simple observation demonstrates a fundamental shortcoming of Metelli's model as a theory of perceptual transparency. Yet neither Metelli nor his followers (e.g., Gerbino, Stultiens, Troost, & de Weert, 1990) have reconciled his formal model with this observation, despite its critical impact on his theory.

In what follows, we argue that one of the main reasons that this puzzle has persisted is because in neither Metelli's experiments nor subsequent investigations has there been an attempt to measure separately how the visual system quantitatively assigns transmittance and lightness to transparent layers. Indeed, although Metelli's model presents quantitative expressions for these two distinct surface attributes, previous studies have typically used experimental tasks that do not allow for separate measurements of these two attributes. In the experiments reported herein, observers made separate matches to either the transmittance or the lightness of a transparent layer. The results of these experiments reveal systematic deviations from the predictions of Metelli's model in both transmittance and lightness. These data provide new insights into how the human visual system computes transparency and motivate a perceptually based theory of transparency that resolves the inadequacy of the episcotister model. In particular, we show that the visual system uses relative Michelson contrast to initiate percepts of transparency and to assign quantitative degrees of opacity to transparent surfaces. More generally, our results reveal the dangers of identifying perceptual theories with the inverse of physical (or generative) models and demonstrate the need to determine experimentally the critical image variables that the visual system uses to compute surface structure.

The Episcotister Model of Transparency

Perhaps the most influential theory of perceptual transparency is due to Metelli (1970, 1974a, 1974b, 1985; Metelli, Da Pos, & Cavedon, 1985). In this section, we review Metelli's episcotister model of transparency and the perceptual predictions he based on this model. We consider both Metelli's original formulation in terms of reflectance values and Gerbino et al.'s (1990) reformulation in terms of luminance values. We then study the extent to which this model provides a general theory of perceptual transparency.

Metelli's Theory

Metelli based his theory of transparency on an episcotister: A disk with reflectance t and an open sector of relative area α is rotated in front of a bipartite background whose two halves **A** and **B** have reflectances a and b (see Figure 1a). When this rotation is sufficiently fast, it leads to the percept of a transparent layer overlying a bipartite background—similar to looking through the blades of a rotating fan (see Figure 1b). The color mixing in the region where the episcotister rotates over the background is given by Talbot's law:

$$p = \alpha a + (1 - \alpha)t. \tag{1}$$

$$q = \alpha b + (1 - \alpha)t. \tag{2}$$

In other words, this situation is treated as one in which two achromatic colors defined by reflectances a and t (or b and t) are mixed in relative proportions α and $1 - \alpha$, respectively, giving rise to a single "fusion color" with reflectance p (or q). Equations 1 and 2 can then be solved for α and t, yielding

$$\alpha = \frac{p - q}{a - b}, \tag{3}$$

[2] In a footnote, Metelli (1974a) attributed this observation to Tudor-Hart (1928).

$$t = \frac{aq - bp}{a + q - b - p}. \tag{4}$$

Metelli argued that perceptual scission in transparency is quantitatively the inverse of color fusion, and hence Equations 1 and 2 describe both color fusion and perceptual scission. In other words, he argued that the visual system is given reflectance values a, b, p, and q, and it computes values of α and t (the transmittance and reflectance of the transparent layer) that satisfy Equations 1 and 2—namely, the solutions in Equations 3 and 4 (Metelli, 1985, p. 187).

Because α in the solution in Equation 3 can only take values between 0 and 1, Metelli's equations predict that the following relations must hold in any image generated by the episcotister setup:

1. Condition 1: $p - q$ must have the same sign as $a - b$ (because $\alpha \geq 0$).
2. Condition 2: The absolute value $|p - q|$ must not exceed the absolute value $|a - b|$ (because $\alpha \leq 1$).

Because Metelli took Equations 1 and 2 to describe perceptual transparency, he argued that these relations (Conditions 1 and 2) can be taken as image conditions that must be satisfied in order for the central region in Figure 1b (i.e., areas **P** and **Q**) to be seen as transparent. Condition 1 says that if region **A** is darker than region **B**, then region **P** must be darker than region **Q**, and vice versa. In other words, the two halves of the central region (i.e., areas **P** and **Q**) must preserve contrast polarity relative to the two halves of the surrounding region (i.e., areas **A** and **B**). Condition 2 says that the central region must have a smaller reflectance difference than the surround. Both of these conditions are clearly satisfied in Figure 1b. Violations of these conditions are shown in Figure 2. In Figure 2a, region **A** is darker than region **B**, but region **P** is lighter than region **Q**—and this disrupts the percept of transparency. In Figure 2b, the reflectance difference within the disk is greater than outside—and this again suppresses the scission of the central disk into multiple layers. (Note, however, that the surrounding region may now appear transparent.) Hence, at least to a first-order approximation, these two constraints seem to predict when the central region in these displays will be seen as transparent. We consider in more detail the perceptual validity of these constraints in the section When to Scission: I. Luminance Conditions for Transparency.

In addition to these conditions on reflectance values, Metelli (1974a, 1974b) and Kanizsa (1979) also pointed to the role of figural conditions in the perception of transparency. Broadly, these may be classified into two kinds: The first requires continuity of the contour on the underlying bipartite surface, while the second requires continuity of the boundary of the putative transparent layer—at the locations where these two sets of contours intersect. Disruptions in the continuity of the underlying contour are shown in Figure 3b; these destroy the percept of a transparent disk overlying a black-and-white background. Disruptions in the continuity of the transparent layer are shown in Figure 3, c and d. In the extreme, if the two halves in the center of Figure 3a are completely separated, they no longer appear transparent (see Figure 3c; see also Metelli, 1974b). Similarly, if these two halves are shifted vertically relative to each other, resulting in discontinuities on the boundary of the putative filter (see Figure 3d), the percept of transparency is again weakened. These effects are also obtained if the disruptions in continuity involve tangent and higher order discontinuities (Kanizsa, 1979), especially if such discontinuities constitute boundaries between perceived parts (Singh & Hoffman, 1998).

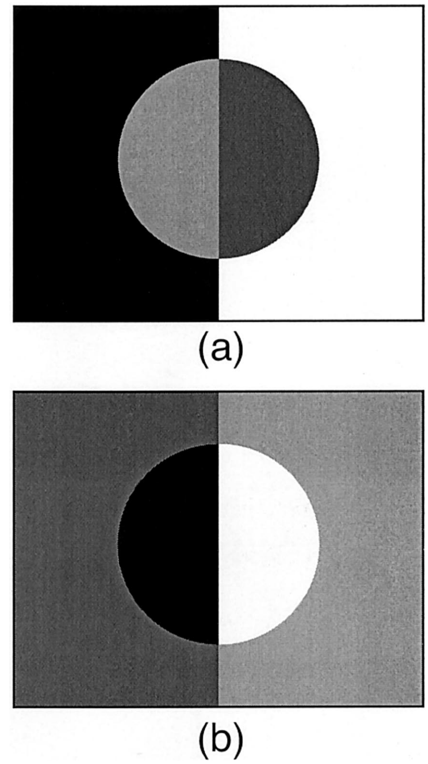

Figure 2. Violations of Metelli's qualitative constraints. (a) Violation of the polarity constraint: Region **A** (left, outside disk) is darker than **B** (right, outside disk), but region **P** (left half-disk) is lighter than **Q** (right half-disk). As a result, the central disk no longer appears transparent. (b) Violation of the magnitude constraint: The reflectance difference within the central region is greater than that in the surround. Again, this suppresses the percept of transparency in the central region. (Note, however, that the surrounding region may now appear transparent.)

Luminance Version of Metelli's Equations

One peculiar aspect of Metelli's equations is that they are written entirely in terms of reflectance values. Because these equations are taken to describe both color fusion and perceptual scission into separate layers, this has the unintuitive implication that reflectance is the variable that is perceptually split into multiple contributions. However, the visual system is only "given" luminance values, not reflectance values. Reflectance is a surface attribute and, as with other surface attributes, it must be inferred by the visual system based on the pattern of luminance values in the image. Hence a perceptual theory of transparency must begin with luminance values rather than reflectance values.

Gerbino et al. (1990) showed that Metelli's Equations 1 and 2 can also be written in terms of luminance values. Moreover, solutions formally identical to Equations 3 and 4 can be derived in the luminance domain under the assumption of uniform illumination. If the illuminant has intensity I, the light P reaching the eyes

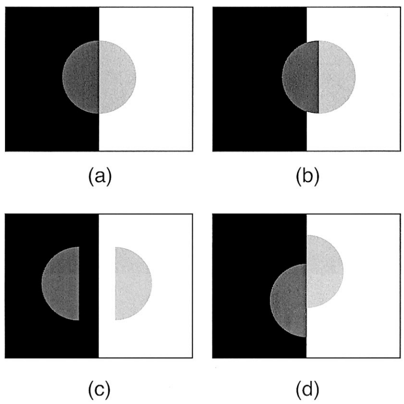

Figure 3. Figural conditions for transparency by Metelli (1974a, 1974b) and Kanizsa (1979). (a) Figural conditions are optimal. (b) The contour dividing the bipartite background must not undergo discontinuous jumps at locations where it meets the boundary of the putative transparent layer. (c) The two gray regions must unite into a coherent surface. (d) The boundary of the transparent layer must not have discontinuities at locations where it meets the underlying contour.

from region **P** (where the episcotister rotates over the background area with reflectance a) is the sum of two components:

$$P = \alpha aI + (1 - \alpha)tI. \tag{5}$$

The first component αaI is the light from the background surface that reaches the eyes after passing through the episcotister (see Figure 4). The second component $(1 - \alpha)tI$ is the light that is reflected directly by the rotating episcotister. (Note that tI is the luminance that would be projected by the opaque portion of a stationary episcotister; see Figure 1a.) Setting $A = aI$ and $T = tI$, we obtain the analogue of Equation 1 in the luminance domain, namely, $P = \alpha A + (1 - \alpha)T$. Similarly, for region **Q** we have $Q = \alpha B + (1 - \alpha)T$. These two equations then yield solutions formally identical to Equations 3 and 4:

$$\alpha = \frac{P - Q}{A - B}, \tag{6}$$

$$T = \frac{AQ - BP}{A + Q - B - P}. \tag{7}$$

The term α in the solution in Equation 6 is the same as in the solution in Equation 3, and its restriction to values between 0 and 1 gives conditions similar to Conditions 1 and 2 above, but in terms of luminance values.

The above analysis models the situation in which the transparent layer and the underlying layer receive the same amount of illumination (see Figure 4). Modified versions of these equations may also be written in cases in which the two layers fail to be equally illuminated. This may occur, for instance, if the transparent layer lies very close to the background surface—so that light from the illuminant is able to reach the background surface only after being attenuated by the transparent layer. Moreover, in this case, repetitive reflections between transparent layer and the background surface may also become significant (Beck, Prazdny, & Ivry, 1984). The modified equations for these cases are considered in Appendix A.

Although Equations 5–7 were derived from an episcotister model, the same equations are also obtained from a generative model that consists of a screen (i.e., mesh) placed in front of a bipartite background. In this case, the transparent layer is modeled as a material surface containing a large number of holes that are too small to be resolved individually (Richards & Stevens, 1979). The transmittance α of the screen corresponds to the areal density of the holes, and the term t now refers to the reflectance of the material portion of the screen. Hence in Equation 5, αaI corresponds to the light that is projected from the background surface after passing through the screen, and $(1 - \alpha)tI$ is the light reflected directly by the screen (see Figure 4). The term tI is multiplied by $1 - \alpha$ because this is the proportion of the screen that is occupied by material surface. It should be noted that there is an ambiguity in the usage of the term *reflectance* when referring to a transparent layer. It could refer to the reflectance of the material portion of the transparent layer (t), such as the reflectance of the dust particles in a dust cloud or the reflectance of the material that forms a screen. Alternatively, it could refer to the reflectance of the transparent layer as a whole (*effective reflectance;* Gerbino et al., 1990), that is, the total proportion of light reflected back from a dust cloud or a screen. Denoting this latter quantity by f, the two are related simply by $(1 - \alpha)t = f$. We use the term *reflectance* in the former sense, because this usage is consistent with the term t in Metelli's equations.

Summary of the Episcotister Model

In the context of four-intensity displays such as in Figure 1b, Metelli and Gerbino et al. (1990) thus provided the following answers to the two basic questions concerning perceptual transparency:

1. The central region (consisting of regions **P** and **Q**) is seen as transparent when the **P**–**Q** border has the same contrast polarity as the **A**–**B** border (the *polarity constraint*), and the **P**–**Q** border has

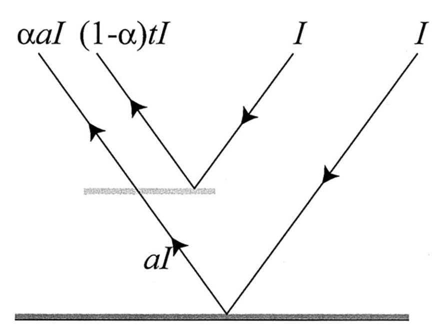

Figure 4. Generative model of transparency based on Gerbino et al. (1990). In the context of an episcotister rotating over a bipartite background, this model yields equations that are formally identical to Metelli's equations, but in the luminance domain. I is the illuminant intensity, a is the reflectance of the background surface, and α is the transmittance of the partially transmissive surface. As a result, αaI is the light from the background surface that reaches the eyes after passing through the partially transmissive surface. (The light aI reflected from the background surface is attenuated by a factor of α.) $(1 - \alpha)tI$ is the light that is reflected directly by the partially transmissive surface.

a lower luminance difference than the **A**–**B** border (the *magnitude constraint*).

2. When the central region is seen as transparent, the transmittance and luminance attributed to the transparent layer are given by the solutions in Equations 6 and 7.

However, both the qualitative constraints for initiating percepts of transparency (expressed as the above polarity and magnitude constraints) and the quantitative predictions for perceived transmittance and lightness (expressed in the solutions in Equations 6 and 7) require the assumption that the transparent layer is homogeneous in transmittance and reflectance, that is, α and t are constant over the transparent layer. (Recall that α and t have to be identical in Equations 1 and 2 in order for the solutions in Equations 3 and 4 to follow.) This assumption has been termed *balanced transparency* (Metelli, 1974a, 1974b; Metelli et al., 1985). This means that naturally occurring forms of transparency involving smoke or fog are outside of the domain of Metelli's theory, because these typically lead to percepts of unbalanced transparency.

Furthermore, Metelli's model of transparency assumes that perceptual decomposition in transparency is consistent with the generative equations of color fusion. However, as we have seen, Metelli's (1974a) own observation—that a black episcotister looks more transmissive than a white episcotister with the same open sector—suggests that there is something wrong with this assumption. Moreover, there is also recent evidence from the chromatic domain to suggest a dissociation between perceptual transparency and physical transparency. D'Zmura, Colantoni, Knoblauch, and Laget (1997), for example, created transparency displays using equiluminous color shifts. These displays generate the percept of transparency even though such shifts are impossible to generate with a physically realizable filter. Thus, one needs to test directly what image properties the visual system uses to initiate percepts of transparency and how it assigns surface attributes to transparent layers.

Extending the Generative Model

As we noted above, Metelli's model makes specific predictions concerning both when to scission and how to scission—but it does so only in the context of four-intensity displays, such as in Figure 1b. This is due to two assumptions that are built into the generative model—one pertains to the transparent layer and the other to the underlying surface. First, the model assumes balanced transparency, that is, that the transparent layer is homogeneous in transmittance and reflectance. Indeed, Metelli et al. (1985) explicitly stated that this model makes no predictions regarding unbalanced forms of transparency. Second, it assumes that the background surface consists of two distinct-reflectance regions (with uniform reflectances a and b). Consistent with these restrictions, the study of perceptual transparency has largely been restricted to displays containing four distinct intensities.

In natural viewing conditions, however, the transparent layer is usually not homogeneous in transmittance and reflectance, and the background surfaces are often textured—containing a continuous range of intensities rather than a small number of discrete values. (Consider, e.g., viewing natural scenes through fog or haze.) In this section, we develop a simple extension of Metelli's equations that allows for both of these possibilities. This permits us to test the perceptual validity of these generative equations using a richer, more articulated class of displays. This has practical significance because, in pilot work, we found that independent judgments of surface opacity and lightness are more reliable when the region of transparency is defined by a continuous range of luminance values—rather than by two subregions of constant luminance. In addition, the extension provides a framework within which unbalanced forms of transparency can be studied. For example, the analysis reveals that a minimal assumption under which qualitative constraints and quantitative solutions can be derived is that of "locally balanced" transparency. In other words, rather than assuming that the transparent surface has constant transmittance and reflectance, we need only assume that changes in these attributes are smooth and gradual (see Figure 5 for examples of unbalanced transparency).

In general, then, the transmittance and reflectance of a screen or filter are functions of position—$\alpha(x, y)$ and $t(x, y)$, respectively—and the background surface is defined by a distribution of reflectance values $a(x, y)$ (Anderson, 1999). Assuming that the intensity of the illuminants on the background surface and transparent layer are $I_a(x, y)$ and $I_t(x, y)$, respectively, this gives the following expression for the distribution of light intensities in the image:

$$L(x, y) = \alpha(x, y)a(x, y)I_a(x, y) + [1 - \alpha(x, y)]t(x, y)I_t(x, y). \quad (8)$$

Equation 8 reveals how unconstrained the problem of determining surface transmittance and reflectance is in general: The distribution of light intensities $L(x, y)$ given to the visual system is consistent with infinitely many distinct combinations of distributions for $a(x, y)$, $\alpha(x, y)$, $t(x, y)$, $I_a(x, y)$, and $I_t(x, y)$. The luminance version of Metelli's model is a special case that derives its solutions under three assumptions: (a) The transparent layer is homogeneous in transmittance and reflectance, (b) the background contains only two reflectances, and (c) the illumination is uniform.

By setting $A(x, y) = a(x, y)I_a(x, y)$ and $T(x, y) = t(x, y)I_t(x, y)$, Equation 8 can be rewritten purely in terms of luminance values:

$$L(x, y) = \alpha(x, y)A(x, y) + [1 - \alpha(x, y)]T(x, y). \quad (9)$$

Equation 9 says that at each image location (x, y), the image luminance $L(x, y)$ is linearly related to $A(x, y)$, where $A(x, y) = a(x, y)I_a(x, y)$ is the distribution of light intensities that is obtained when the background surface is seen in plain view (i.e., without the intervening transparent surface). If $\alpha(x, y)$ and $T(x, y)$ are completely unconstrained, then no solutions can be derived because the

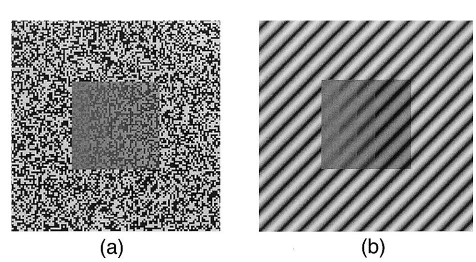

Figure 5. Examples of unbalanced transparency. The transparent layer in (a) is seen as having inhomogeneous transmittance—increasing from left to right. The transparent layer in (b) is unbalanced in both transmittance and lightness. Going from left to right, it becomes darker and more transmissive.

slope and intercept of this linear function are free to vary arbitrarily from point to point. In order to derive local solutions, we make the assumption that the transparency is locally balanced. In other words, we assume that α and T vary in a smooth and gradual manner, so that they can be treated as being effectively constant over local image neighborhoods (e.g., see Figure 5). Discontinuities in α and T are permitted as long as there are sufficiently large regions of smooth variation in between, over which mean luminance and luminance range can be reliably computed.

Over such a local neighborhood, the following solutions can be derived (see Appendix B):

$$\alpha = \frac{L_{\max} - L_{\min}}{A_{\max} - A_{\min}} = \frac{L_{\mathrm{range}}}{A_{\mathrm{range}}}, \tag{10}$$

$$T = \frac{A_{\mathrm{range}}\, L_{\mathrm{mean}} - A_{\mathrm{mean}}\, L_{\mathrm{range}}}{A_{\mathrm{range}} - L_{\mathrm{range}}}. \tag{11}$$

These solutions are simple generalizations of the solutions to Metelli's equations in their luminance version—namely, the solutions in Equations 6 and 7. In particular, the local solution for α is given by the ratio of luminance ranges when the background surface is viewed through the transparent layer to when it is seen unobscured. The solution for T gives the luminance assigned locally to the transparent layer. The local surface reflectance t can be computed if further assumptions can be made about the illuminant(s).

In the next section, we report two matching experiments that investigate how the visual system actually assigns opacity and lightness to transparent surfaces. The solutions derived above allow us to assess the extent to which these perceived surface attributes are predicted by Metelli's generative equations. Later, in the section When to Scission: I. Luminance Conditions for Transparency, we test the predictions of these equations in initiating percepts of transparency.

Matching Experiments: Assigning Surface Attributes

How does the visual system assign opacity and lightness to a transparent surface, given that it has decomposed an image region into two separate layers? The natural strategy for addressing this question is to use stimuli that evoke a vivid percept of distinct layers in depth and then ask observers to make separate judgments of opacity and lightness of the transparent layer. To this end, we used a sinusoidal grating as the background surface (see Figure 6) rather than a bipartite background (see Figure 1b). This has the effect that the region of transparency is defined in the image by a continuous profile of luminance values rather than by just two distinct-luminance regions. This greatly enhances the percept of separate layers and, in turn, makes the judgments of surface attributes more robust. Second, we added disparity to the contour that bounds the region of transparency such that it appeared in front of the high-contrast background. This reinforces the construction of two separate layers in depth and generates a clear percept of a small homogeneous transparent filter floating in front of a high-contrast grating. (See the stereo percept obtained by fusing the stereogram in Figure 6.) The manipulation of relative depth also represents a more ecologically valid form of transparency because, in natural environments, transparency typically occurs in conjunction with such depth information.

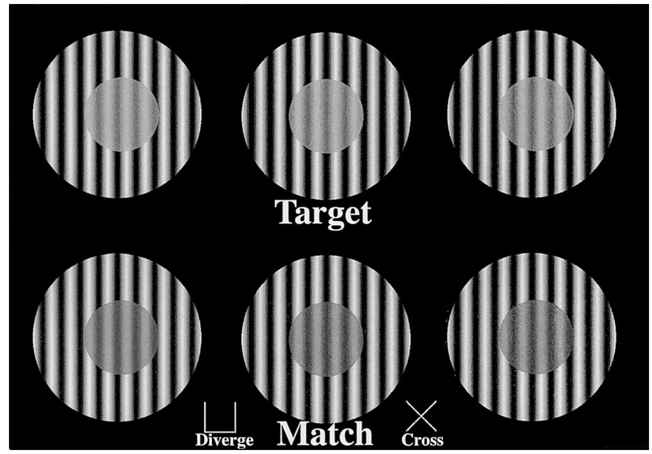

Figure 6. Stimuli used in Experiments 1 and 2. (Cross-fusers should fuse the right two half-images, divergers the left two.) Observers matched either the transmittance (Experiment 1) or the lightness (Experiment 2) of the transparent disk in the target display by adjusting the luminance range (Experiment 1) or mean luminance (Experiment 2) in the central disk of the matching display.

We performed two matching experiments using displays such as in Figure 6. In Experiment 1, observers matched the perceived transmittance of the small target disk by adjusting the luminance range (or amplitude) in the corresponding disk in a matching display. This task required observers to ignore any difference in the lightness of the two transparent disks and match them solely along the dimension of perceived transmittance. In Experiment 2, observers matched the lightness of the transparent disk by adjusting the mean luminance of the matching disk—while ignoring any difference in the perceived transmittance of the two disks.

Experiment 1: Matching Perceived Transmittance

Method

Observers. Three observers with normal or corrected-to-normal vision participated in the experiment. Two were naive to the purposes of the experiment, and the 3rd was an author (Manish Singh).

Stimuli and apparatus. Each test display consisted of a large circular disk containing a vertically oriented sinusoidal grating (see Figure 6). The disk was placed on a black background, and its diameter subtended a visual angle of 4.96°. The spatial frequency of the sinusoidal grating was 1.8 cycles per degree, its mean luminance (A_{mean}) was 44.9 cd/m^2, and its luminance range ($A_{\max} - A_{\min}$) was 84.8 cd/m^2 (contrast = 0.945). A smaller disk was placed inside the large disk (the "target disk"). This disk had a lower contrast but the same mean luminance (L_{mean} = 44.9 cd/m^2). Throughout, we use L to denote luminance values in the region of transparency (here, the small disk) and A to denote luminance values in the nontransparent region (here, the large surrounding disk). The diameter of the smaller disk subtended 2.2°. Four different values of luminance range in the small disk were tested: 10.63, 21.23, 31.83, and 42.44 cd/m^2. These corresponded to Michelson contrast values of 0.118, 0.236, 0.354, and 0.472, respectively. A near disparity of 8.3 min of arc was introduced to the boundary of the small disk. As noted earlier, this generates the percept of a small homogeneous transparent filter floating in front of a high-contrast grating.

A matching display was placed 1.6° below the test display. This was identical to the test display and it contained the same disparity (see Figure

7), but the mean luminance within its small disk varied from trial to trial. Six values of mean luminance were used: 23.7, 30.77, 37.84, 44.91, 51.98, and 59.05 cd/m^2. The luminance range within this small disk was to be adjusted by the observer.

The displays were presented on a Radius PressView 17SR monitor (Radius, Inc., Sunnyvale, CA), which was calibrated so that luminance values were linearly related to the 8-bit look-up-table values (ranging from 0.58 cd/m^2 to 90.14 cd/m^2). All displays were viewed through a mirror stereoscope at an optical distance of ~100 cm.

Procedure. Each observer performed four blocks of trials, one block for each of the four contrast values within the target disk. On each trial, the mean luminance in the small disk in the matching display was randomly set to one of the six preset values. The observer's task was to adjust the luminance range in this small disk using a mouse, such that its perceived transmittance (opacity) matched that of the target disk (see Figure 7). Because the target and matching disks usually differed in their mean luminance, this task required observers to ignore any difference in the lightness of the two transparent disks and match them solely along the dimension of perceived transmittance. The luminance range increased and decreased symmetrically about the preset value of mean luminance. In each block, observers performed 5 adjustments for each of the six values of mean luminance. (Observer R.V.E. performed two such blocks for each of the four contrast values, in order to get the size of the error bars comparable to that of the other 2 observers. This resulted in a total of 10 adjustments per condition.)

Results and Discussion

Observers reported this to be an easy and natural task. The matching data for the 3 observers are shown in Figure 8. Each curve represents an observer's settings of luminance range within the small disk in the matching display to match the perceived transmittance of a fixed target disk. The different curves represent matches to different target disks—with the same mean luminance but different luminance ranges (and hence, different contrasts). Each observer's data thus provide four different iso-transmittance curves.

The most striking property of these data is that the transmittance matches are almost perfectly fit by increasing linear functions that pass through the origin. (See the equations for the linear fits and r^2 values in Table 1.) Thus, as the mean luminance within the matching disk increases, observers set a higher luminance range to match the perceived transmittance of a fixed target disk. Moreover, as the contrast within the target disk increases, so do the slopes of these linear fits.

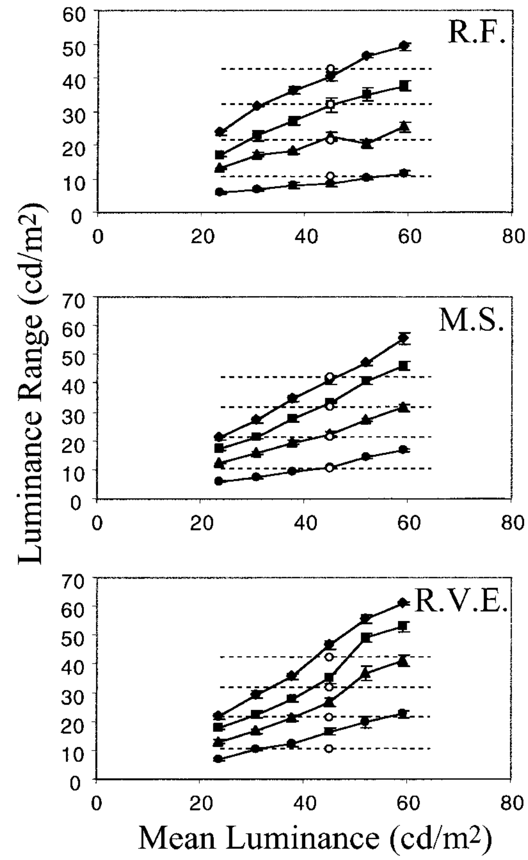

Figure 8. Results of the transmittance-matching task in Experiment 1 for the 3 observers (R.F., M.S., and R.V.E.). The four curves represent transmittance matches to four different target disks with different contrast values (shown in open circles). These iso-transmittance curves deviate systematically from the predictions of Metelli's equations—shown by dashed lines—which predict that the iso-transmittance curves should be flat (i.e., independent of mean luminance). Observers thus systematically overestimate the transmittance of transparent surfaces that darken the underlying surface, and they underestimate the transmittance of transparent surfaces that lighten the underlying surface.

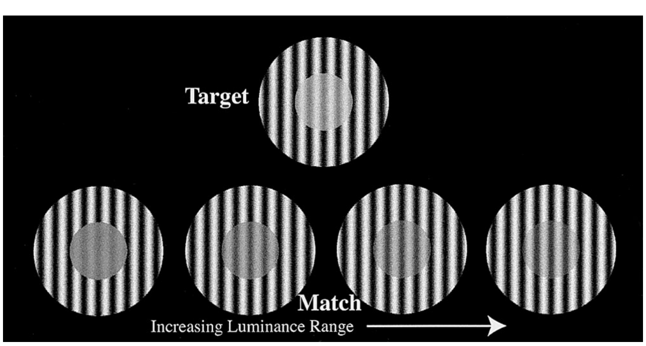

Figure 7. A demonstration of the transmittance-matching task of Experiment 1. Observers matched the perceived transmittance (or degree of opacity) of the transparent disk in the target display by adjusting the luminance range within the central disk of the matching display. Increasing the luminance range makes the transparent disk look more transmissive. Observers were instructed to ignore any difference in lightness between the two transparent disks and base their judgments solely on perceived opacity.

This pattern of results is not at all consistent with the solution (see Equation 10) for α—which is a simple generalization of Metelli's solution (see Equation 6). This solution predicts that the transmittance matches—expressed in terms of luminance range—should be independent of mean luminance. (These predictions are shown by the dashed lines in Figure 8.) This is because the solution for α is given by the ratio of luminance range in the region of transparency (in this case, the central disk) to the luminance range in the unobscured region (in this case, the surround). Because the surrounds are identical in our target and match displays, the transmittance matches should be determined entirely by the luminance range in the region of transparency. Clearly, this prediction is not borne out perceptually.

Table 1
Summary of Model Fits for the 3 Observers in Experiment 1

Test contrast	R.V.E. y	R.V.E. r^2	R.V.E. SE_{slope}	M.S. y	M.S. r^2	M.S. SE_{slope}	R.F. y	R.F. r^2	R.F. SE_{slope}
0.4725	1.021x	.991	.013	0.917x	.995	.011	0.902x	.993	.014
0.3544	0.845x	.964	.016	0.76x	.996	.009	0.684x	.985	.015
0.2364	0.642x	.948	.016	0.52x	.995	.007	0.456x	.976	.013
0.1184	0.363x	.948	.011	0.267x	.993	.004	0.201x	.975	.006

Instead, these data are consistent with the perceived transmittance's being determined by the Michelson contrast in the region of transparency relative to the Michelson contrast in the surrounding region. An easy way to see this is to replot the data in terms of Michelson contrast (see Figure 9), by dividing each luminance range setting by twice the mean luminance in the matching disk:

$$\frac{L_{\max} - L_{\min}}{L_{\max} + L_{\min}} = \frac{L_{\text{range}}}{2 \cdot L_{\text{mean}}}. \quad (12)$$

This requires the assumption that the distribution of luminance values is symmetrical, such that $L_{\text{mean}} = (L_{\max} + L_{\min})/2$, which is true in our displays. As a result of this transformation, each of the data curves becomes essentially independent of mean luminance. This is precisely what one would expect from the matches in Figure 8: Because these take the form $y = kx$, the ratio $y/2x$ ($=k/2$) is independent of x. These results indicate that the critical variable that the visual system uses to assign surface transmittance is Michelson contrast, not luminance range. The results are thus consistent with a modified formula for α (recall the solutions in Equations 6 and 10), in which the numerator is the Michelson contrast in the region of transparency and the denominator is the Michelson contrast in the surround.

Our result that perceived transmittance is determined by relative Michelson contrast—rather than relative luminance difference or range—resolves the long-standing puzzle that originated with Metelli's (1974a) observation that a black episcotister looks more transmissive than a white episcotister with the same physical transmittance (i.e., with the same size sector cut out). Metelli realized that there was something wrong with the perceptual predictions derived from his equations—because his solution for α would predict the two episcotisters to appear equally transmissive. However, he never reconciled this observation with his formal model. The ratio-of-contrasts rule for α, on the other hand, provides a simple and complete account of this observation, because the black episcotister generates a higher Michelson contrast than the white one.

Furthermore, our analysis reveals that the order relation between perceived transmittance and physical transmittance depends on whether the transparent surface decreases (i.e., darkens) or increases (i.e., lightens) the mean luminance projected by the background surface. Observers systematically overestimate the transmittance of darkening transparent surfaces, and they systematically underestimate the transmittance of lightening transparent surfaces. Thus, the perceived transmittance of a transparent surface depends not only on its own physical attributes but also on those of the surface it partially occludes. Moreover, the magnitude of this overestimation (or underestimation) of surface transmittance is directly proportional to the extent to which the transparent surface decreases (or increases) the mean luminance of the underlying surface.

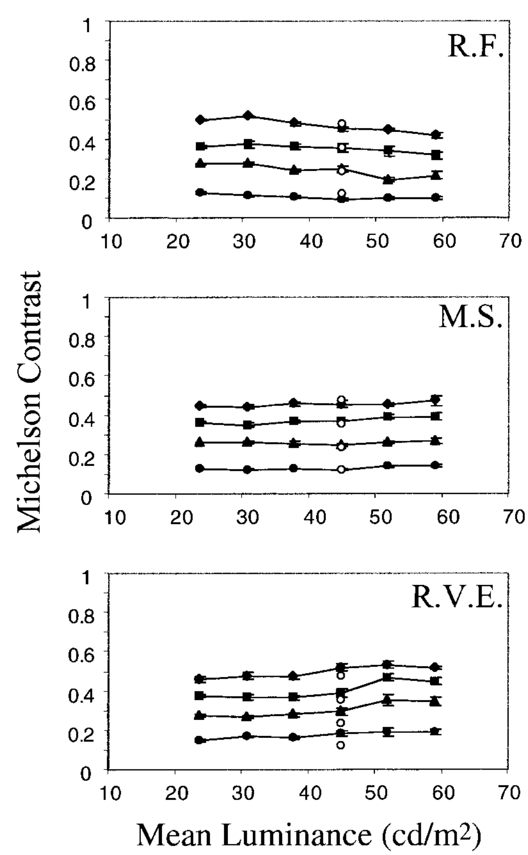

Figure 9. Results of Experiment 1 replotted in terms of Michelson contrast. The four curves represent transmittance matches to four different target disks with different contrast values (shown in open circles). Results are shown for the 3 observers (R.F., M.S., and R.V.E.). Each data curve flattens out as a result of the transformation from luminance range to Michelson contrast. This indicates that the visual system uses the Michelson contrast in the region of transparency—and not luminance difference or luminance range—to assign perceived transmittance to transparent surfaces.

Experiment 2: Matching Surface Lightness

Method

Observers. The same 3 observers participated as in Experiment 1.

Stimuli and apparatus. The apparatus and stimulus dimensions were the same as in Experiment 1. As before, the large disk in both the test and the matching display remained fixed throughout the experiment (A_{mean} = 44.9 cd/m^2; A_{range} = 84.8 cd/m^2). This time, however, the luminance range within the small target disk remained fixed (L_{range} = 31.83 cd/m^2), and five different values of mean luminance were tested: 27.24, 36.07, 44.9, 53.74, and 62.58 cd/m^2. In the matching display, the luminance range with the central disk varied from trial to trial. Six values of luminance range were used: 0, 10.61, 21.22, 31.83, 42.44, and 53.06 cd/m^2. These correspond to Michelson contrast values of 0, 0.118, 0.236, 0.355, 0.473, and 0.591, respectively. The mean luminance in the central disk of the matching display was to be adjusted by the observer.

Procedure. Each observer performed five blocks of trials, one block for each of the five values of mean luminance in the target disk. On each trial, the luminance range in the small disk of the matching display was randomly set to one of the six preset values. The observers' task was to adjust the mean luminance in this matching disk using a mouse, so that its lightness (i.e., the lightness of the perceived transparent disk) matched that of the target disk (see Figure 10). Any difference in the perceived opacities of the target and matching disks was to be ignored. In each block, observers performed 5 adjustments for each of the six values of luminance range. (As in Experiment 1, Observer R.V.E. performed two such blocks for each of the five values of mean luminance, for a total of 10 adjustments per condition.)

Results and Discussion

Observers reported the lightness-matching task to be much less natural and substantially more difficult. The matching data for the 3 observers are shown in Figure 11. Each curve represents an observer's settings of mean luminance within the small matching disk to match the lightness of a fixed target disk. The different curves represent matches to different target disks—with the same luminance range but different mean luminances. Each observer's data thus provide five different iso-lightness curves.

Unlike the transmittance matches of Experiment 1, the lightness matches exhibit a fair degree of variability across observers. The matches of Observer M.S. exhibit some degree of convergence toward the mean luminance of the surround, with increasing luminance range within the matching disk. Observer R.F. exhibits this convergence only weakly for the left-most and right-most curves, whereas Observer R.V.E. does not show this effect.

In order to better understand these results, it is instructive to compare them with three possible theoretical predictions. First, consider the predictions of the solution (see Equation 11) for T—which is a simple extension of Metelli's solution (see Equation 7). Denoting variables pertaining to the match display by primed symbols (T', A', L') and those of the target display by unprimed symbols (T, A, L), we obtain the predicted settings of mean

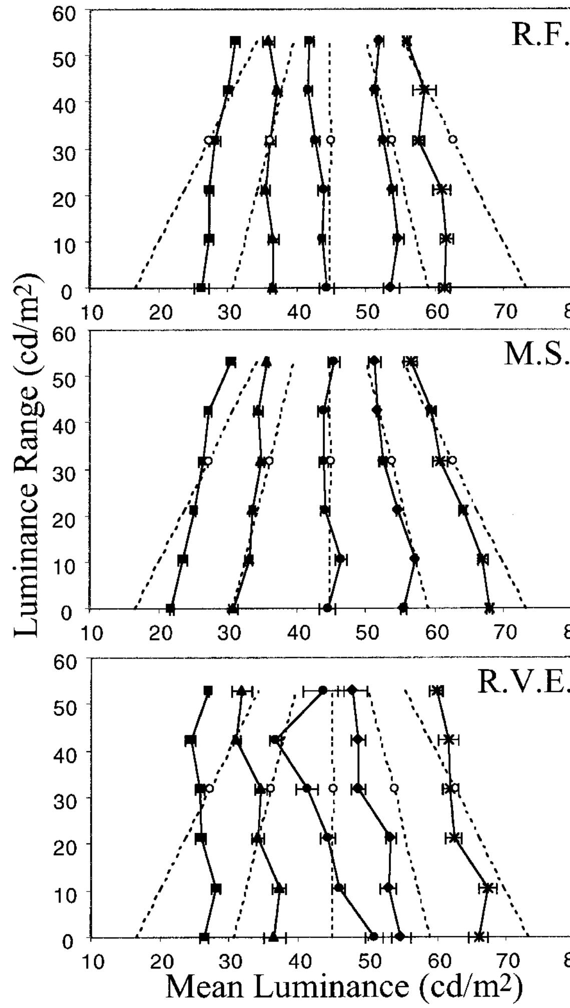

Figure 11. Results of the lightness-matching task in Experiment 2 for the 3 observers (R.F., M.S., and R.V.E.). The five curves represent lightness matches to five different target disks with different values of mean luminance (shown in open circles). Although the iso-lightness curves for Observers R.F. and M.S. show some trend toward convergence, the extent of this convergence is much weaker than predicted by the term T (see Equations 7 and 11) in Metelli's equations. (These predictions are shown by dashed lines.) The lightness attributed to a transparent surface is thus strongly biased toward the mean luminance in the region of transparency and is consistent with only a partial "discounting" of the luminance generated by the underlying layer.

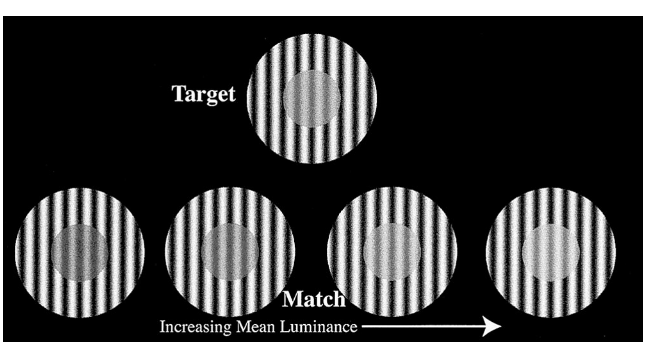

Figure 10. A demonstration of the lightness-matching task of Experiment 2. Observers matched the lightness of the transparent disk in the target display by adjusting the mean luminance within the central disk of the matching display. Increasing the mean luminance makes the transparent disk look lighter. Observers were instructed to ignore any difference in the perceived opacity of the two transparent disks and to base their judgments solely on their lightnesses.

luminance (L'_{mean}) in the match display by equating $T' = T$. Expanding the respective expressions for T and T' using Equation 11, and using the fact that the surrounds are the same in the match and target ($A = A'$), we obtain the following solution for L'_{mean}:

$$L'_{\text{mean}} = L_{\text{mean}} + \left(\frac{A_{\text{mean}} - L_{\text{mean}}}{A_{\text{range}} - L_{\text{range}}}\right)(L'_{\text{range}} - L_{\text{range}}). \quad (13)$$

The iso-lightness curves are thus predicted to be linear functions with slopes $(A_{\text{mean}} - L_{\text{mean}})/(A_{\text{range}} - L_{\text{range}})$. (These are shown by the dashed lines in Figure 11; see also the top panel of Figure 12.) Because the luminance range within the surround of the target display is always higher than that within the central disk ($A_{\text{range}} - L_{\text{range}} > 0$), the sign of the slope is determined entirely by the sign of the numerator $A_{\text{mean}} - L_{\text{mean}}$. Thus, the slopes of the iso-lightness curves are predicted to be positive when the target disk is darker than its surround ($L_{\text{mean}} < A_{\text{mean}}$), negative when the target disk is lighter than its surround ($L_{\text{mean}} > A_{\text{mean}}$), and zero when the center and surround have the same mean luminance ($L_{\text{mean}} = A_{\text{mean}}$). Moreover, the magnitudes of these slopes are such that they predict complete convergence as the luminance range within the match disk approaches the luminance range of the surround (i.e., $L'_{\text{mean}} \to A_{\text{mean}}$, as $L'_{\text{range}} \to A_{\text{range}}$). In the limit, this corresponds to the case in which the transparent layer becomes fully transmissive and hence does not contribute to the image. In sum, the matches predicted by Metelli's solution for T are linear functions that pass through the respective target points and converge to the mean luminance of the surround as the luminance range within the matching disk increases.

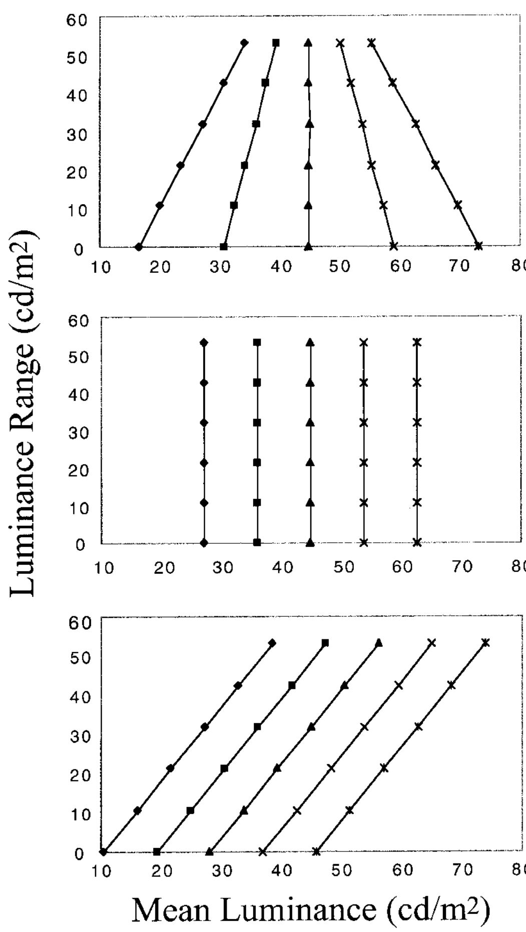

Figure 12. Three possible theoretical predictions for the lightness-matching task of Experiment 2: matches predicted by the term T in Metelli's equations (top; see Equations 7 and 11) matches predicted by the local mean luminance in the central disk (middle); and matches predicted by "effective luminance" F (Gerbino et al., 1990)—the additive term in Metelli's equations, $F = (1-\alpha)T$ (bottom). The matching data (in Figure 11) are largely a compromise between the predictions in the top and middle panel, but are not captured by the prediction in the bottom panel.

A second theoretical prediction (plotted in the middle panel of Figure 12) is one that would be expected if the central region did not scission perceptually into two separate layers and the luminance within it were thus attributed entirely to the target disk (i.e., none of the luminance was attributed to the underlying surface). In this case, there would thus be no "discounting" of any luminance from the underlying layer, and lightness matches would be predicted simply by the mean luminance within the target disk:

$$L'_{\text{mean}} = L_{\text{mean}}. \quad (14)$$

The graphs in Figure 11 demonstrate that observers' lightness matches are largely a compromise between these two predictions. Although there is a slight convergence of the data curves with increasing luminance range, there is also a strong bias toward the vertical, relative to the predictions based on Metelli's solution for T (these are shown superimposed in dashed lines). Even for Observer M.S., who exhibited the strongest convergence, the extent of convergence is considerably weaker than that predicted by the solution for T. (Table 2 shows the slopes of the best linear fits to the lightness-matching data and the slopes predicted by the solution for T.) This trend toward the vertical indicates that when observers attempt to estimate the lightness of a transparent layer, their matches are biased strongly toward the mean luminance in the region of transparency. Thus observers are able to separate and "discount" only incompletely the contributions of the underlying surface as they attempt to judge the lightness of the transparent layer. These matches can be described by the following expression:

$$L'_{\text{mean}} = L_{\text{mean}} + K\left(\frac{A_{\text{mean}} - L_{\text{mean}}}{A_{\text{range}} - L_{\text{range}}}\right)(L'_{\text{range}} - L_{\text{range}}) \quad (0 < K < 1), \quad (15)$$

where K is a parameter that modulates the extent of scission, which is reflected in the degree of convergence. $K = 0$ corresponds to the case of no convergence (Equation 14), and $K = 1$ corresponds to the case of the full convergence predicted by the term T (Equation 13). This degree of convergence provides a measure of the extent to which observers are able to discount the contribution of the underlying layer. It is therefore not surprising that Observer M.S. shows the strongest convergence because this observer is more practiced than the other 2 in making judgments on transparent layers.

A final theoretical prediction we consider is one proposed by Gerbino et al. (1990), based on "effective luminance," that is, the product $F = (1 - \alpha)T$. This is essentially the additive term in

Table 2
Predicted and Observed Slopes for the 3 Observers in Experiment 2

Test mean luminance	Slope predicted by Metelli[a]	Observer R.F. Slope	R.F. *SE*	M.S. Slope	M.S. *SE*	R.V.E. Slope	R.V.E. *SE*
62.58	−0.333	−0.110	.047	−0.222	.033	−0.129	.055
53.74	−0.167	−0.054	.031	−0.105	.047	−0.142	.057
44.91	0	−0.053	.029	−0.01	.039	−0.182	.081
36.07	0.167	−0.004	.029	0.079	.024	−0.112	.049
27.24	0.333	0.089	.031	0.15	.024	−0.023	.031

[a] Metelli, 1974b. Predictions were derived by plugging the values used in the current experiment into Metelli's equations.

Metelli's equation ($P = \alpha A + F$; recall Equation 5). Gerbino et al. proposed that the lightness of transparent layers is matched based on effective luminance F (see the Relation to Previous Empirical Work on Transparency section for a discussion of Gerbino et al.'s, 1990, experiment). Within the context of our generative setup involving textured backgrounds, the solution for F may be written as follows:

$$F = L_{\mathrm{mean}} - \left(\frac{L_{\mathrm{range}}}{A_{\mathrm{range}}}\right) A_{\mathrm{mean}}. \tag{16}$$

Setting $F = F'$, and using the fact that the surrounds are identical in the target and match displays, we obtain the following lightness matches predicted by effective luminance:

$$L'_{\mathrm{mean}} = L_{\mathrm{mean}} + \left(\frac{A_{\mathrm{mean}}}{A_{\mathrm{range}}}\right)(L'_{\mathrm{range}} - L_{\mathrm{range}}). \tag{17}$$

The main characteristic of these predicted matches is that their slopes are positive and constant throughout (equal to $A_{\mathrm{mean}}/A_{\mathrm{range}}$), irrespective of the mean luminance L_{mean} of the target disk. The iso-lightness curves based on these predictions are shown in the bottom panel of Figure 12. As is evident from these curves, effective luminance fails to capture even the qualitative trend in the lightness-matching data (see Figure 11). As we noted above, the term T does capture this qualitative trend somewhat—although, quantitatively, it predicts a stronger convergence than is actually exhibited in observers' data.

The relatively high variability across observers is consistent with their reports that lightness matching is a considerably more difficult task than transmittance matching and is, as a result, more prone to strategic influences on the part of the observers. The increased difficulty in estimating the lightness of a transparent surface is not surprising given the greater number of variables that must be taken into account in making this computation. Whereas the computation of surface transmittance requires taking into account only two contrast values (in the transparent and unperturbed image regions, respectively), the computation of lightness of a transparent layer involves taking into account two values of mean luminance plus an estimate of surface transmittance (which involves the two contrast values). Nevertheless, to the extent that observers can perform this task, their lightness matches are not well predicted by the solution derived from Metelli's equations.

When to Scission: I. Luminance Conditions for Transparency

The preceding experiments investigated how the visual system assigns attributes to transparent surfaces once perceptual decomposition into multiple surfaces has occurred, but they did not directly address the issue of when the visual system initiates such decomposition. As we noted earlier, in the context of four-intensity displays generated by the episcotister model (recall Figure 1b), Metelli proposed two qualitative constraints to predict when the central region will be seen as transparent: The two halves of the central region must have the same contrast polarity as the two halves of the surround (the polarity constraint), and the two halves of the central region must have a lower luminance difference (the magnitude constraint).[3]

The role of the polarity constraint in initiating percepts of transparency has been fairly uncontroversial. In expressing the importance of this constraint, some researchers have classified X junctions on the basis of the ordinal relations that can occur between the four luminance values. Beck and Ivry (1988) noted that drawing line segments successively in the order of increasing lightness values leads to either Z configurations, C configurations, or criss-cross configurations (see Figure 13). They suggested that junction type provides an important cue to perceptual transparency. For example, in the left display (Z configuration), either of the two squares can be seen as transparent; in the middle display (C configuration), only the lower square can be seen as transparent; and in the right display (criss-cross configuration), neither square is seen as transparent. Adelson and Anandan (1990) directly used the polarity constraint to classify X junctions into nonreversing, single reversing, and double reversing. In a nonreversing junction, both of the edges that form the X junction preserve contrast polarity (left display); in a single-reversing junction, only one of the two edges preserves contrast polarity (middle display); and in a double-reversing junction, neither edge preserves contrast polarity (right display).

Although the above classification schemes capture the polarity constraint, they do not take into account Metelli's magnitude constraint (Condition 2). Anderson (1997) proposed a qualitative

[3] As before, we consider the luminance version of Metelli's equations.

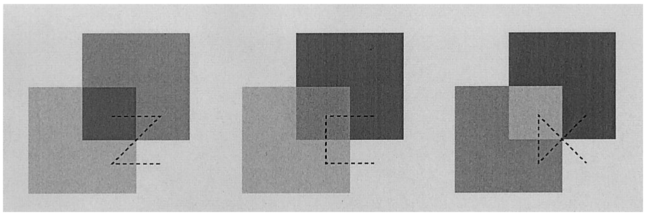

Figure 13. Junction classification schemes based on Metelli's polarity constraint (Adelson & Anandan, 1990; Beck & Ivry, 1988). The ordinal relations between the four luminance values are predictive of whether both (left; nonreversing junctions/*Z*-configuration), one (middle; single-reversing junctions/*C*-configuration), or neither (right; double-reversing junctions/criss-cross configuration) of the two squares can be seen as transparent.

rule that articulates a sufficient local image condition to initiate perceptual scission: "When two aligned contours undergo a discontinuous change in contrast magnitude, but preserve contrast polarity, the lower contrast region is decomposed into two causal layers" (p. 420). Figure 14a demonstrates this rule. Figure 14a shows a transparency display, along with one of its T junctions shown enlarged on the right. The two aligned vertical segments that form the "top" of this T junction preserve contrast polarity (the left side is darker for both segments), and the top half has lower contrast. As a result, the top half scissions into two distinct layers, and the display on the left is seen as a light-colored filter overlying a black-and-white background.

More recently, Anderson (1999) extended his analysis to displays that do not contain discontinuous changes in contrast. Given an image with graded variations in contrast, the visual system is faced with the problem of determining which image regions correspond to surfaces that are being seen in plain view and which—if any—correspond to surfaces perturbed by an intervening transparent layer. Anderson (1999) proposed an anchoring principle that states that human vision treats the regions of highest contrast in an image as unobscured (in other words, with full transmittance; $\alpha = 1$) and regions of lower contrast as obscured by an overlying transparent layer.

A critical question both for the qualitative rule and for the anchoring principle is the following: How are "lower contrast" and "higher contrast" to be defined in order to predict percepts of transparency? According to Metelli's magnitude constraint, the relevant notion of *contrast* is defined in terms of luminance differences. The displays that Anderson (1997, 1999) studied all involved polarity preserving T junctions (see Figure 14a), or "fuzzy" T junctions (whose "stems" are defined by graded variations in luminance). For these junction types, the notion of lower contrast is defined unambiguously because the luminance value remains unchanged on one side of the aligned contours that preserve contrast polarity. For example, in the T junction in Figure 14a, both the gray and the black on the left are compared with the same white on the right—and the gray–white border unambiguously has lower contrast. Hence, in that context, it was sufficient to define contrast as the size of the luminance difference across a contour. (See Appendix C for a proof that luminance difference and Michelson contrast make the same predictions of lower contrast in the case of polarity preserving T junctions.)

Note, however, that T junctions are somewhat special in the context of transparency because they are generated only in those situations in which the transparent layer projects the same luminance as one of the two sides of an underlying contrast edge. For example, Figure 14a is consistent only with a white transparent layer overlying the black-and-white stripes. If the transparent layer were any other color, the image junctions obtained would be X junctions (see also Watanabe & Cavanagh, 1993b). Hence, in general, situations involving transparency are more likely to generate X junctions in the projected image, rather than T junctions.

In the context of X junctions, which of two sides is assigned lower contrast depends critically on the precise definition of contrast used. In particular, luminance difference (or range) and Michelson contrast can make opposing predictions of what constitutes lower contrast. For example, in the nonreversing X junction shown schematically in Figure 14b, the region to the left of the vertical contour has a lower luminance difference (i.e., $X - Y < Z - W$), whereas the region on the right has a lower Michelson contrast—that is, $[(X - Y)/(X + Y)] > [(Z - W)/(Z + W)]$.

In this section, we address the following question: What quantity must be lowered in an image region in order for the percept of transparency to be initiated? We first consider the predictions that follow from the episcotister model and then generalize these to the

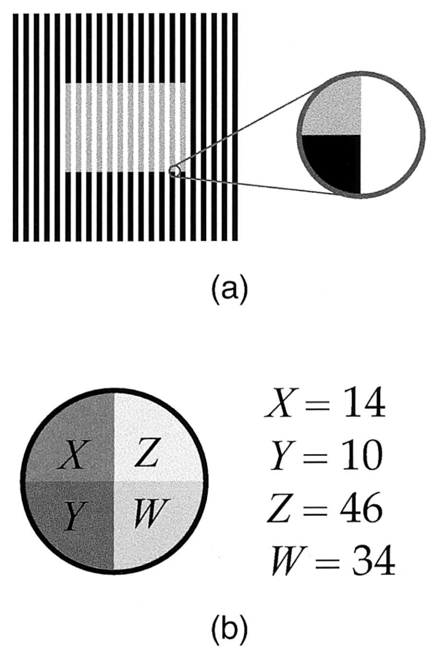

Figure 14. (a) An example demonstrating Anderson's (1997) qualitative rule for initiating perceptual scission into multiple layers. The two aligned vertical contours of the enlarged T junction preserve contrast polarity, and the upper half has lower contrast. As a result, the upper half scissions into two layers: a light-colored transparent layer and an underlying layer with black and white stripes. Note that, in the case of polarity preserving T junctions, the notion of *lower contrast* is defined unambiguously. (b) For X junctions, on the other hand, luminance difference and Michelson contrast can make opposing predictions of what constitutes "lower contrast." For the luminance values shown (*X, Y, Z, W*), the left side of the X junction has a lower luminance difference, but the right side has a lower Michelson contrast.

extended generative equations. Recall the solutions that follow from the episcotister model (i.e., Equations 6 and 7). Because the restriction $\alpha \geq 0$ on the solution in Equation 6 forces the **P–Q** border to have the same contrast polarity as the **A–B** border, we may assume without loss of generality that $A > B$ and $P > Q$. The restriction $\alpha \leq 1$ on the solution in Equation 6 then implies that $P - Q \leq A - B$. In other words, the region of transparency necessarily has a lower luminance difference than the surrounding region.

Note that the two qualitative constraints that Metelli derived are both based on restrictions on the solution in Equation 6 for α. However, the solution in Equation 7 for T is also restricted. Although Metelli recognized this, he thought that "the other formula . . . is more complicated and does not offer the opportunity of deriving simple predictions" (Metelli, 1974a, p. 105). This may have been in part because Metelli thought of the terms in these solutions as reflectance values. Beck et al. (1984) similarly concluded that qualitative constraints based on the solution for T are perceptually meaningless, but for a different reason. They argued that the terms in Metelli's equations should be construed as lightness values, and as a result, the expression for T involves sums and products of lightness values.

Construing A, B, P, and Q as luminance values, however, we demonstrate that a simple and meaningful constraint can be derived from restrictions on the solution in Equation 7 for T. To see this, note that the numerator in the solution in Equation 7 can be rewritten as follows:

$$AQ - BP = \frac{1}{2}[(P+Q)(A-B) - (A+B)(P-Q)]$$

$$= \frac{1}{2}(A+B)(P+Q)\left(\frac{A-B}{A+B} - \frac{P-Q}{P+Q}\right). \qquad (18)$$

Because T is a luminance value, the natural restriction on the solution in Equation 7 is $T \geq 0$. Using Equation 18, this restriction can be rewritten as follows:

$$\frac{1}{2}(A+B)(P+Q)\,\frac{\left(\dfrac{A-B}{A+B} - \dfrac{P-Q}{P+Q}\right)}{(A-B)-(P-Q)} \geq 0. \qquad (19)$$

Because $(A + B)$ and $(P + Q)$ are always nonnegative (being sums of luminance values), the restriction $T \geq 0$ is thus equivalent to

$$\frac{\left(\dfrac{A-B}{A+B} - \dfrac{P-Q}{P+Q}\right)}{(A-B)-(P-Q)} \geq 0. \qquad (20)$$

Note that the numerator in this expression is the difference in Michelson contrasts between the unobscured surround and the transparent center, whereas the denominator is the difference in their luminance ranges. For this expression to be positive, the numerator and denominator must either both be positive or both be negative. However, as we have seen earlier, the restriction $\alpha \leq 1$ already forces this denominator to be positive, namely, $(A - B) - (P - Q) \geq 0$. Hence the restriction $T \geq 0$ now forces $[(A - B)/(A + B)] - [(P - Q)/(P + Q)] \geq 0$. In other words, Metelli's equations predict that both luminance difference and Michelson contrast must be lowered in the region of transparency.[4] (This conclusion easily generalizes to the extended version of Metelli's equations that we considered in Extending the Generative Model; see Appendix D.) Thus, even though Metelli did not realize this, Michelson contrast does enter into the predictions that follow from his equations. However, it enters only into the qualitative conditions for initiating scission—not into the expression for transmittance α (Equation 6).

If perceptual scission qua transparency is consistent with the above prediction, one would predict that the visual system initiates such scission in those image regions where both luminance range and Michelson contrast are lowered relative to adjoining regions. In order to plot this prediction, consider the diagram in Figure 15b. Let the unobscured region (the surround in Figure 15a) be defined by mean luminance L_0, and luminance range R_0, and the region of putative transparency (the center in Figure 15a) be defined by mean luminance L and luminance range R. The region under the horizontal dashed line then gives the set of (L, R) combinations for which the luminance range is less than (L_0, R_0). If lower contrast for transparency is defined based on lowering luminance range, this would be the set of (L, R) values for which the central region appears transparent. Similarly, the region under the oblique dashed line gives the set of (L, R) combinations for which the Michelson contrast is less than (L_0, R_0). If the perception of transparency is initiated based on lowering Michelson contrast, then this would be the set of permissible (L, R) values for which the central region appears transparent. Finally, both luminance range and Michelson contrast are lowered in the region shaded in gray (this is the intersection of the first two regions). If perceptual scission is initiated only when luminance range and Michelson contrast are both lowered, then this region would define the set of permissible (L, R) values for which the central region appears transparent.

Note that the horizontal dashed line and the oblique dashed line divide the (L, R) plane into four regimes. These are labeled $R+C+$, $R+C-$, $R-C+$, and $R-C-$, depending on whether the luminance range (R) and Michelson contrast (C) are increased (+) or decreased (−) relative to the region defined by (L_0, R_0)—that is, the unobscured surround in Figure 15a. Hence $R+C+$ and $R-C-$ denote the regimes where luminance range and Michelson contrast make the same predictions of what constitutes lower contrast, and $R+C-$ and $R-C+$ denote the regimes where they make opposing predictions. Note that if the mean luminance of the center (L) is the same as that of the surround (L_0), then luminance range and Michelson contrast always make the same predictions. Hence the predictions of luminance range and Michelson contrast are best separated when the center and surround differ sufficiently in their mean luminances.

Experiment 3: What Constitutes Lower Contrast for Transparency?

Figure 16 shows two stereo displays similar to Figure 6, in which near disparity has been given to the boundary of the central

[4] An alternative proof of this is the following: $(P - Q)/(P + Q) = [\alpha(A - B)]/[\alpha(A + B) + 2(1 - \alpha)T] \leq [\alpha(A - B)]/[\alpha(A + B)] = (A - B)/(A + B)$. The inequality follows because $\alpha \leq 1$ and $T \geq 0$, so that $2(1 - \alpha)T \geq 0$. The proof in the main text has the advantage, however, of demonstrating how Michelson contrast enters into the expression for T.

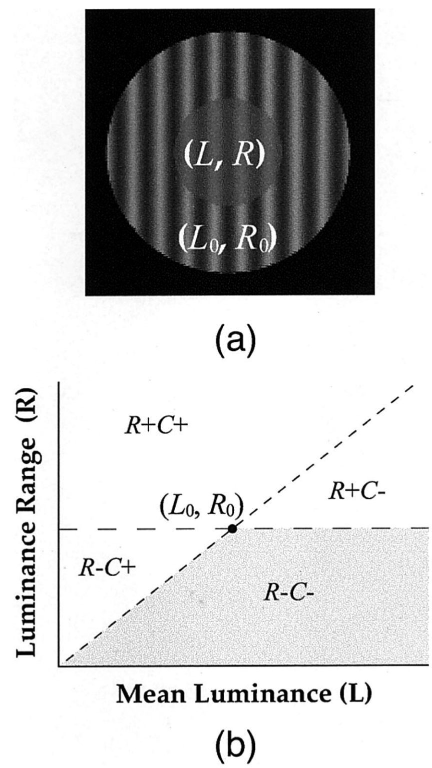

Figure 15. Demonstrating the predictions of perceived transparency based on luminance range and Michelson contrast. Let the surround in display (a) be defined by mean luminance L_0 and luminance range R_0. Display (b) plots the regimes (demarcated by dashed lines) in which luminance range (R) and Michelson contrast (C) are increased (+) or decreased (−), relative to the surround. For example, $R-C+$ denotes the regime in which luminance range is lowered but Michelson contrast is increased relative to the surround. The region shaded in gray ($R-C-$) denotes the regime in which both luminance range and Michelson contrast are lowered.

disk, relative to that of the large surrounding disk. Although this places the boundary of the central disk closer in depth, it does not by itself specify which of the two sides of this boundary (the center or the surround) will get pulled to that depth. In Figure 16a, the contrast in the central disk is unambiguously lower than in the large surrounding disk and, as in Figure 6, this generates the percept of a small transparent disk floating in front of a sinusoidal background. In Figure 16b, on the other hand, the contrast in the central disk has been made unambiguously higher than the surround, and this changes the percept dramatically: Now the surrounding annulus decomposes into two separate layers, and the central region looks like a hole in a transparent layer, through which the underlying surface is seen in plain view. Thus contrast information plays a key role, above and beyond stereoscopic depth, in determining perceived surface structure. At what point does the above switch in perceived surface structure take place? For the case depicted in Figure 16, this question has a clear answer because the center and surround have the same mean luminance. As a result, the central region actually becomes indistinguishable from the background when its luminance range is made equal to that of the surround. However, this is not true more generally, when the center and surround have different mean luminances. What, in general, is the set of permissible values of luminance range that the central region can take and still appear transparent?

Method

Observers. The same 3 observers participated as in Experiments 1 and 2.

Stimuli and apparatus. The apparatus and stimulus dimensions were the same as in Experiments 1 and 2. However, only one stereo pair was presented on the computer monitor, because there were no separate test and match displays in this experiment. The large surrounding disk remained unchanged throughout the experiment (A_{mean} = 37.84 cd/m^2; A_{range} = 45.26 cd/m^2). The mean luminance within the central disk varied from trial to trial. Six values of mean luminance were used: 20.17, 27.24, 34.31, 41.38, 48.44, and 55.51 cd/m^2. The luminance range within the small disk was to be adjusted by the observer using a mouse.

Procedure. Each observer performed two blocks of trials—one for each of two different tasks. For the first task, observers were asked to indicate the highest luminance range within the central disk that clearly generated the percept of a small transparent disk overlying the large sinusoidal-grating background. For the second task, observers were asked to indicate the lowest luminance range within the central disk that did not generate a percept of transparency in the small disk. The rationale behind these two tasks was to identify regions of (L, R) space where transparency is clearly seen within the central disk, regions where transparency is clearly not seen within the central disk, and regions where the percept is ambiguous (the "zone of confusion"). On each trial, the mean luminance in the small disk was randomly set to one of the six preset values, and the observer adjusted its luminance range in accordance with the task. For each task, observers performed five adjustments for each of the six values of mean luminance.

Results

The data for the 3 observers are shown in Figure 17. The data exhibit a marked asymmetry between darkening filters that de-

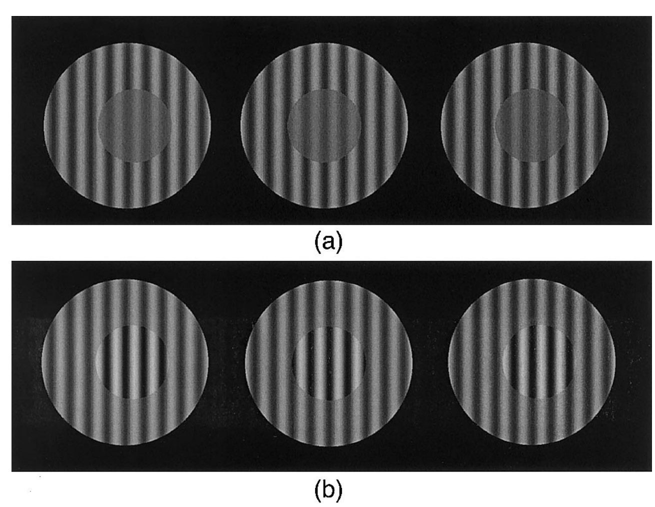

Figure 16. (a) A stereogram in which the central region has unambiguously lower contrast than the surround. As in Figure 6, this generates the percept of a small transparent disk overlying the sinusoidal grating. (b) If the contrast in the central region is made unambiguously higher than the surround, it is now the surround that scissions into two separate layers, whereas the central region is seen as a hole in the transparent layer through which the underlying grating is seen in plain view.

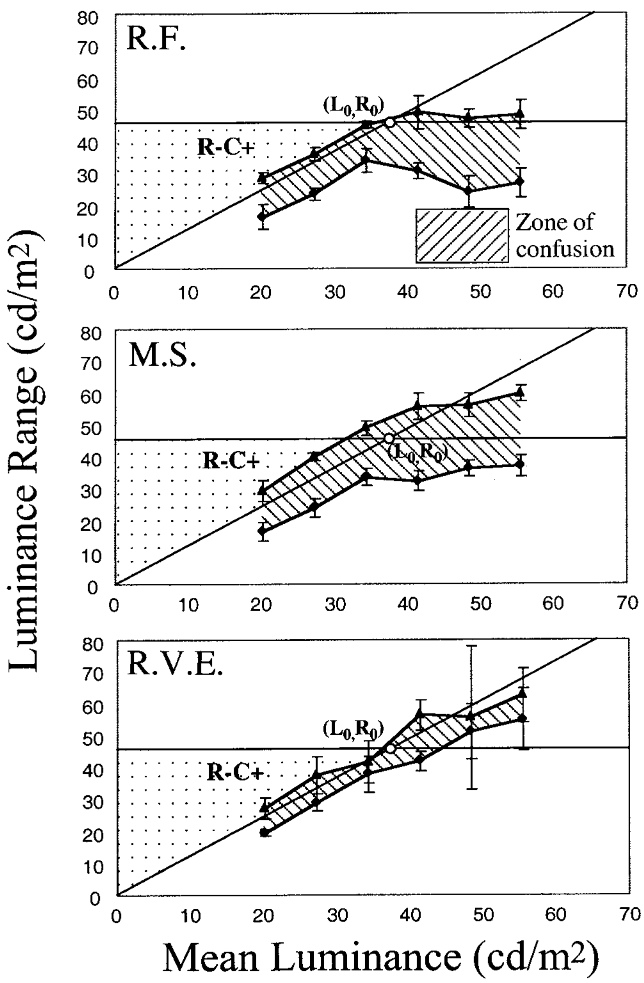

Figure 17. Results of Experiment 3 for the 3 observers (R.F., M.S., and R.V.E.). The lower curve corresponds to the settings for the first task, namely, to indicate the highest luminance range that clearly generates the percept of transparency in the central region. The upper curve corresponds to the settings for the second task, namely, to indicate the lowest luminance range that clearly does not generate the percept of transparency in the central region. The area in between gives a "zone of confusion" in which observers are unsure as to whether or not the central region contains transparency. Note that observers do not perceive transparency in the regime $R-C+$ (which represents decreased luminance range and increased Michelson contrast) even though luminance range is lowered relative to the surround—thus providing a counter to Metelli's magnitude constraint. L_0 and R_0 represent the mean luminance and luminance range, respectively, within the target disk. The error bars represent standard errors of the mean.

crease the mean luminance of the underlying surface—that is, all data points to the left of the point (L_0, R_0)—and lightening filters that increase the mean luminance of the underlying surface—that is, all data points to the right of (L_0, R_0). The 3 observers agree closely for darkening filters, that is, when the mean luminance of the central disk is lower than that of the surround. Here, observers' data curves for the two tasks follow closely the oblique-line prediction based on Michelson contrast. In other words, observers see transparency when the Michelson contrast within the central disk is lower than in the surround and fail to see transparency otherwise. This provides a clear counter to Metelli's prediction that transparency is seen when the luminance differences (or, more generally, luminance range) is lowered relative to the surround—recall Metelli's magnitude constraint—because observers do not see the central region as transparent in the regime $R-C+$ even though the luminance range is lowered.

For lightening filters, that is, when the mean luminance of the central disk is higher than that of the surround, observers reported that the decision as to whether or not the central region is transparent was considerably more difficult. This difficulty is also reflected in observers' data in that the data curves for the two tasks are either considerably more separated—implying a larger zone of confusion (for Observers M.S. and R.F.; recall that responses on these two tasks were based on either clearly seeing, or clearly not seeing, transparency in the central region)—or else the data exhibit considerably larger error bars (for Observer R.V.E.). Moreover, in this regime, there is also more variability across observers. Whereas the data curves of Observers R.F. and M.S. follow the horizontal-line prediction based on luminance range (albeit with large zones of confusion), those of observer R.V.E. are closer to the oblique-line prediction based on Michelson contrast.

Two main conclusions can be drawn from these results. First, there is an asymmetry between transparent layers that darken the underlying surface and those that lighten the underlying surface—such that judgments of transparency are easy and reliable for darkening transparent layers but difficult and more variable for lightening layers. It has often been suggested that visual mechanisms for computing transparency may have been derived evolutionarily from mechanisms for computing shadows—given that shadows are more prevalent in the natural environment. Although this remains a somewhat controversial issue, the asymmetry observed in our results is at least consistent with such a claim—given that shadows can only decrease mean luminance.

Second—and more important from the point of view of testing Metelli's model—the case of darkening filters provides a clear contradiction to Metelli's magnitude constraint for predicting perceived transparency, namely, that luminance range must be lowered in the region of putative transparency. In particular, observers consistently failed to see transparency in the regime $R-C+$ even though luminance range was lower in the central region relative to the surround, and their data curves followed closely the oblique prediction line based on lowering Michelson contrast. This suggests that, as in the case of assigning perceived transmittance, Michelson contrast is a critical variable that the visual system uses to initiate the perceptual construction of transparency.

When to Scission: II. Conditions of Figural Grouping

As we noted earlier (recall Figure 3), in addition to luminance constraints, certain conditions of figural grouping must also be satisfied in order to initiate percepts of transparency. Indeed, figural grouping can be such a strong cue that, in some cases, a strong global configuration consisting of two overlapping figures can sometimes lead naive observers to make reports of transparency—even if the local luminance relations do not support transparency (Beck et al., 1984). For example, in Figure 18a, the grouping of the respective boundaries of two figures is so compelling that one is clearly aware of a square and a disk overlapping; and this in itself may lead some naive observers to sometimes

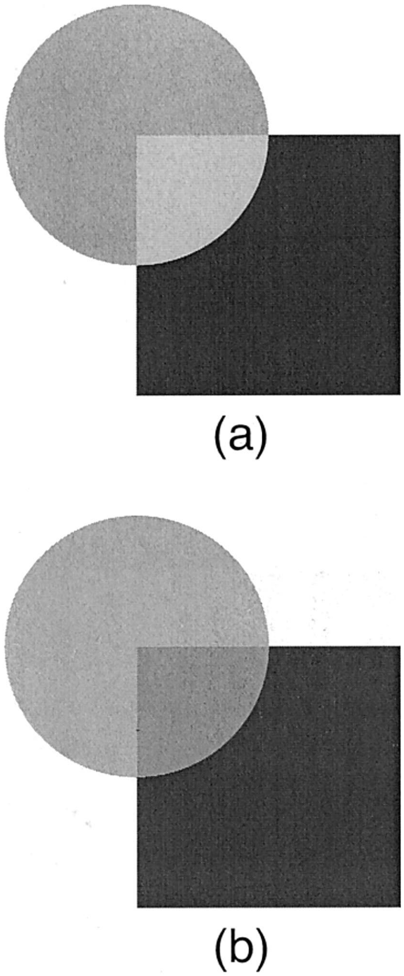

Figure 18. (a) Strong figural grouping and global configuration can lead some naive observers to sometimes make reports of transparency, even if contrast polarities are reversed (Beck et al., 1984; Beck & Ivry, 1988). Note, however, that one does not perceive two different lightnesses in the region of overlap in the display in (a)—one surface lightness seen through another—the way that one does in the display in (b). This phenomenological difference has also been shown to have measurable consequences for speeded recognition (Watanabe & Cavanagh, 1993a) and negative priming (Loula et al., 2000). Our focus is on transparency that involves a decomposition of local image intensity into multiple surfaces in depth.

make reports of transparency. Note, however, that there is something qualitatively different about the "transparency" seen in such displays. In particular, whereas one is clearly aware of two distinct surfaces in the region of overlap in Figure 18b—a dark surface seen through a gray one—one does not similarly perceive two different lightnesses in the region of overlap in Figure 18a. In other words, whereas the display in Figure 18b involves a scission of local image intensity into two distinct lightnesses, the display in Figure 18a does not.[5] In this article, we are concerned only with transparency that involves such local scission of image intensity.

Conditions of figural grouping for transparency have typically been proposed in the context of images that contain a small number of distinct-luminance regions separated by sharp contours and junctions. For example, in the context of four-intensity displays (see Figures 1 and 3), Metelli (1974b) and Kanizsa (1979) proposed two kinds of conditions: one requiring the continuity of the contour that divides the bipartite background (Figure 3b) and the other requiring the continuity of the contour that bounds the putative transparent layer (Figure 3, c and d). Among these, Anderson's (1997) qualitative rule assigns special importance to the alignment of contours on the background surface.

In the most general situation, however, the underlying surfaces may not contain clearly defined contours. This could occur either because the "contours" on the underlying surface are defined by graded variations in luminance or, more generally, because the underlying surfaces are textured (see Figure 19a). In these cases, the notion of "contour alignment" must be extended to some notion of textural continuity or grouping. The main intuition here is that the display must be consistent with a generic view of the underlying surface continuing behind a transparent layer; for example, in the left of Figure 19a, the textures in the center and surround are continuous with each other, and this display is thus consistent with a generic view of the textured background continuing behind a transparent layer. The qualification of a generic view is important because, according to the principle of *genericity* (or nonaccidentalness) for scene interpretation, the visual system rejects unstable (or accidental) interpretations of image data (Albert & Hoffman, 1995; Biederman, 1987; Freeman, 1994; Koenderink & van Doorn, 1979; Lowe, 1985; Nakayama & Shimojo, 1992; Witkin & Tenenbaum, 1983). Specifically, in the context of transparency, viewing conditions that involve nongeneric or accidental alignments of the transparent layer with underlying contours can prohibit the percept of transparency (see, e.g., Nakayama & Shimojo, 1992; Singh & Hoffman, 1998). For example, observers do not see transparency on the left of Figure 19b even though this display is consistent with a gray transparent layer overlying a black-and-white background. The interpretation of transparency is unstable in this display, however, because small displacements of the transparent layer relative to the underlying surface can lead to large qualitative changes in the junction structure of the image (see the right of Figure 15b). The display on the right, on the other hand, depicts a generic view of a gray transparent layer overlying the same background—and it readily evokes the percept of transparency.

One way to define the notion of textural continuity between two adjoining textures is by requiring that the only variables that distinguish them at their shared boundary be their contrast and mean luminance. In displays such as in Figure 19a, this is tanta-

[5] This phenomenological difference has also been demonstrated to have measurable consequences for speeded recognition and negative priming. With brief presentations (60–150 ms), Watanabe and Cavanagh (1993a) found that participants were significantly more accurate in naming two overlapping digits if the region of overlap preserved contrast polarity (analogous to Figure 14b), than if it did not (analogous to Figure 18a). More recently, Loula, Kourtzi, and Shiffrar (2000) used a negative priming paradigm (e.g., Tipper, 1985) in which observers were asked to match one of two overlapping shapes. On some trials, the unattended shape on trial n became the attended shape on trial $n + 1$. Loula et al. found that negative priming occurred when the region of overlap of the two shapes reversed contrast polarity (so that the two shapes were more difficult to segment), and positive priming occurred when the region of overlap preserved polarity.

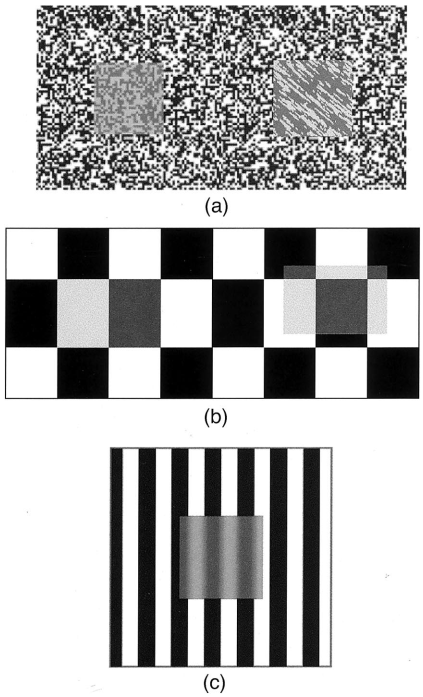

Figure 19. (a) Extending figural conditions for transparency to textured displays: The texture in the low-contrast image region must be continuous with the texture in the high-contrast image region. This ensures that the image is consistent with a generic view of the underlying surface continuing behind a transparent layer. (b) The role of the *principle of genericity* (or nonaccidentalness) in the perception of transparency. The left display is physically consistent with a transparent layer overlying a checkered background. However, it involves nongeneric (or accidental) alignments of the transparent layer with the contours on the background and does not evoke the percept of transparency. If the transparent layer is displaced slightly to yield a generic placement, it is readily seen as transparent. (c) Translucent surfaces scatter light as well and thus blur the contours of underlying surfaces seen through them. A complete account of figural grouping must take into account such blurring effects as well.

mount to saying that the boundary between the two regions should become imperceptible if the contrast of the low-contrast region is uniformly increased to match that of the high-contrast region (in a way that does not reverse local contrast polarities). This captures the intuition that although transparent layers change the luminance and contrast across underlying contours and textures, they do not change these contours and textures themselves. The right side of Figure 19a shows a display in which this definition is not satisfied. Here the central region has been replaced with a texture patch that does not group with the surround. As a result, this region is much less likely to scission into two separate layers.

In those cases in which textures contain clearly defined contours, two textures are continuous by the above criterion only if their contours are aligned at their mutual border. Textural continuity is greatly strengthened, moreover, by the presence of polarity-preserving contours that continue across the shared border between two textured regions—even if these contours exist only at low spatial frequencies. The random-dot textures in the left display of Figure 19a and the gratings in Figure 6, for example, both satisfy the criterion for textural continuity articulated above. However, the continuity of the oriented elements of the gratings across the contrast border makes for a more compelling sense of textural continuity.

The above analysis also shows that the anchoring principle (which assigns full transmittance to the highest contrast regions in an image) is actually contingent on the presence of textural continuity between the low-contrast and high-contrast regions. In the absence of such continuity (e.g., see the right side of Figure 19a), there is no decomposition into separate layers and hence no notion of an overlying transparent layer (or perturbation) that can be assigned varying degrees of transmittance.

A caveat concerning the preceding analysis is that it assumes that transparent layers do not alter the contours and textures on underlying surfaces. This assumption is valid for models of transparency that have typically been studied—namely, those involving only the light-transmitting properties of transparent layers but ignoring their light-scattering properties. However, if one takes into account the light-scattering properties of translucent media (e.g., Blinn, 1982; Hanrahan & Krueger, 1993; Kubelka, 1954; Singh & Anderson, 2002), then the above definition becomes too restrictive. A translucent layer that scatters light in addition to absorbing some proportion of it also makes the contours and textures on the underlying surface appear more fuzzy. An example is shown in Figure 19c, where placing a translucent layer over a high-contrast square-wave grating not only lowers its contrast but also blurs its edges. Observers still perceive transparency in this display (i.e., they see two separate layers in the central region) even though the textures in the center and surround are not continuous by the strict criterion articulated above. (In particular, making the contrast in the center equal to that in the surround does not make their shared boundary imperceptible.) This suggests that the criterion articulated above should be weakened by allowing for the blurring effects of translucent filters. In particular, the figural criterion for initiating scission must be extended by postulating that textural continuity modulo blurring (in the region of lower contrast) is sufficient to initiate perceptual scission in the region of lower contrast (see Singh & Anderson, 2002).

Relationship to Lightness, Brightness, and Apparent Contrast

The reader may have noted that monocular versions of the displays we used in our experiments have been used previously in the vision literature—however, emphasizing a very different aspect of these displays, namely, apparent contrast. Indeed, our choice of displays was motivated in part by the belief that these effects of apparent contrast are closely related to visual mechanisms that construct the percept of transparency. In this section, we articulate a relationship of perceptual transparency to phenomena of apparent contrast, as well as to the perception of lightness and brightness.

In a series of experiments, Chubb, Sperling, and Solomon (1989) demonstrated that a fixed texture patch appears to have lower contrast when it is surrounded by high-contrast texture than when it is surrounded by low-contrast texture (see Figure 20). They termed this effect *contrast–contrast,* in analogy with the well-known effect of simultaneous contrast, wherein a uniform gray patch looks darker when it is surrounded by a light-colored region than when it is surrounded by a dark-colored region. In

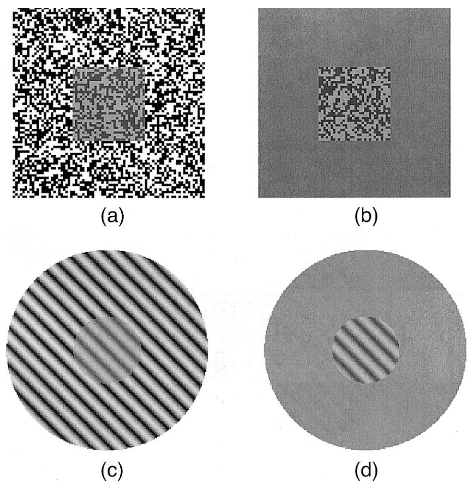

Figure 20. A demonstration of contrast–contrast (Chubb et al., 1989; Solomon et al., 1993). The contrast in a texture patch appears to be lower when it is surrounded by a high-contrast texture, as shown in (a) and (c), than when it is surrounded by a low-contrast texture as shown in (b) and (d). Note, however, that apparent contrast is not the only perceptual dimension along which these display pairs differ. In particular, the displays on the left appear to have a transparent surface overlying the background textures, whereas the displays on the right do not. From "Texture Interactions Determine Perceived Contrast," by C. Chubb, G. Sperling, & J. Solomon, 1989, *Proceedings of the National Academy of Sciences USA, 86,* p. 9632, Figure 1. Copyright 1989 by the National Academy of Sciences. Adapted with permission.

order to account for this phenomenon, Chubb et al. (see also Solomon, Sperling, & Chubb, 1993) postulated the existence of low-level mechanisms of lateral inhibition, similar to those used to account for simultaneous contrast, that modulate the apparent contrast of the texture patch. In other words, the output gain of a texture-sensitive unit is normalized by the rectified sum of the outputs of surrounding units. Hence, a unit's response is suppressed if it is surrounded by units that are highly activated—resulting in a lowering of apparent contrast in a texture patch that is surrounded by high-contrast texture.

There are two basic issues that arise in terming this phenomenon *contrast–contrast.* The first is the description of the phenomenology. In particular, *apparent contrast* is asserted to be the primary phenomenological dimension along which the two central patches differ. This description then motivates a theory that attempts to account for this dimension of phenomenological change, namely, lateral-inhibitory mechanisms of contrast normalization. We ask the following questions: (a) Does the difference in apparent contrast constitute a complete description of the phenomenology associated with these displays? (b) Does lateral inhibition provide a complete account of contrast phenomena?

Consider first the phenomenology associated with the displays in Figure 20. Although a difference in apparent contrast is evident, another perceptual difference between these pairs of displays is that the central regions in Figure 20, a and c, appear as transparent disks overlying high-contrast backgrounds, whereas the central regions in Figure 20, b and d, do not (see also D'Zmura & Singer, 1999). This phenomenological difference is at least as salient as the difference in apparent contrasts. Moreover, this difference persists even if the apparent contrasts in these pairs are matched. For example, one can gradually decrease the physical contrast in the texture patch in Figure 20d and at some point minimize the difference in apparent contrast with the central texture patch in Figure 20c. However, even when this occurs, the two displays are not metamers. The central region in Figure 20c is readily seen as containing two separate layers in depth, but the central region in Figure 20d does not similarly decompose into two separate layers. This difference in perceived structure is thus another dimension of visual phenomenology—one that is not describable solely in terms of apparent contrast.

Consider now the explanation of contrast–contrast in terms of lateral inhibition. In the contrast-modulated version of White's illusion shown in Figure 21a (D'Zmura & Singer, 1999; Spehar & Zaidi, 1997), the four rectangular texture patches on top have the same physical contrast as the four patches at the bottom. Moreover, the patches on top are bordered by more low-contrast texture, whereas the ones at the bottom are bordered by more high-contrast texture. Thus lateral inhibition would predict that the patches at the bottom should appear to have lower contrast. The perceived effect, however, is just the opposite: The patches on top appear to have lower contrast. This suggests that lateral inhibition is (at least) insufficient as an account of contrast phenomena.

The insufficiency of lateral inhibition in the contrast domain is not surprising, given that lateral inhibition has been shown to be similarly incomplete as an explanation of lightness and brightness phenomena. In a large number of studies, configural factors leading to the interpretation of perceived surfaces (Adelson, 1993; Gilchrist, 1977; Hochberg & Beck, 1954) and perceived surface curvature (Knill & Kersten, 1991), as well as relationships between surfaces, such as occlusion (Todorovic, 1997; Zaidi, Spehar, & Shy, 1997) and transparency (Adelson, 2000; Anderson, 1997; Somers & Adelson, 1997), have all been shown to lead to dramatic changes in perceived brightness. Because lateral inhibition, by its very nature, simply performs a kind of image processing on the two-dimensional image representation, it does not take into account surface-level representations that are subsequently constructed by the visual system. Thus, such results are all beyond its explanatory scope.

In many first-order displays, brightness effects that cannot be explained by lateral inhibition are easily understood when considered from the point of view of layered surface representations. For example, Anderson (1997) argued that the standard White's illusion (see Figure 21b; White, 1979) can be understood from the point of view of layered representations. According to this account, the alignment of the vertical contours and the lowering of contrast within the patches cause these patches to undergo perceptual scission into multiple layers. As a result, part of the darkness of the four patches on top is attributed to the underlying black stripes; hence the patches themselves appear lighter. Similarly, part of the lightness of the patches at the bottom is attributed to the underlying white stripes, so that these patches appear darker. Thus, whereas lateral inhibition predicts the opposite sign of the illusion, a scission account predicts not only the correct sign but also certain aspects of the size of the illusion, for example, the fact that the illusion can be stronger than the corresponding simultaneous contrast effect (see Anderson, 1997).

As another example of the role of transparency in modulating brightness, consider the stereo display in Figure 22 (based on

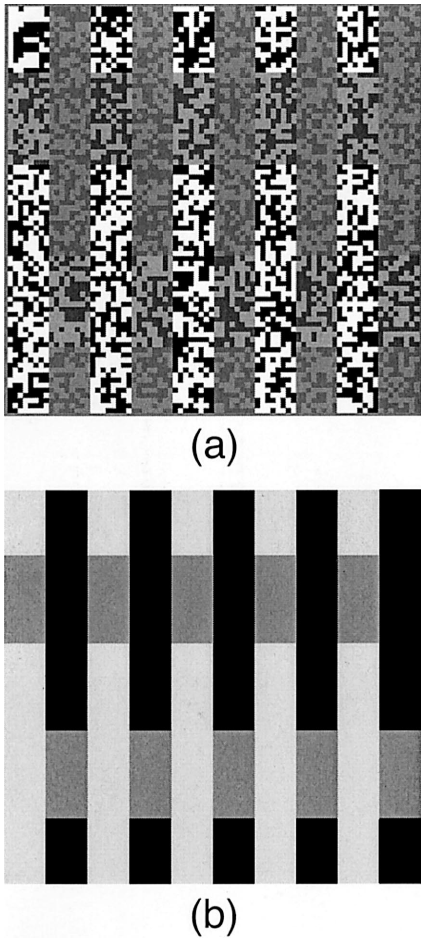

Figure 21. (a) The second-order White's illusion demonstrates the insufficiency of lateral inhibition in predicting apparent contrast (D'Zmura & Singer, 1999; Spehar & Zaidi, 1997). The four texture patches at the top are bordered by more low-contrast texture, whereas the ones at the bottom are bordered by more high-contrast texture. Hence lateral inhibition predicts that the patches at the bottom should appear to have lower contrast. The observed effect is just the reverse: The patches at the top appear to have lower contrast. (b) The standard White's illusion (White, 1979) similarly demonstrates the insufficiency of lateral inhibition in the brightness domain. The four patches at the top appear darker even though they are bordered by more black and should thus appear lighter according to lateral inhibition. Top panel from "New Configurational Effects on Perceived Contrast and Brightness: Second-Order White's Effects," by B. Spehar & Q. Zaidi, 1997, *Perception, 26,* p. 411, Figure 2. Copyright 1997 by Pion Limited, London. Adapted with permission. Bottom panel from "A New Effect of Pattern on Perceived Brightness," by M. White, 1979, *Perception, 8,* p. 413, Figure 1. Copyright 1979 by Pion Limited, London. Adapted with permission.

Nakayama, Shimojo, & Ramachandran, 1990; Varin, 1971). Compare the two percepts that are obtained when the large diamond is given near versus far disparity relative to the four white disks. In the former case, the diamond is seen as a transparent surface that overlies the four disks. In the latter, it is seen as an opaque surface viewed through four portholes. Note that the diamond appears darker when it is seen as a transparent surface than when it is seen as an opaque surface. When it is seen as a transparent surface, part of the lightness of the diamond is attributed to the underlying white disks. This, in turn, causes the diamond itself to appear darker. Of importance, the two images are identical in both cases—only their presentation to the two eyes has been switched. Thus this difference in brightness cannot be attributed to low-level image differences but results solely from differences in relative depth placement of perceived surfaces and their relative opacities.

Given that the standard contrast–contrast displays (e.g., Figure 20, a and c) are consistent with the percept of a transparent surface overlying a high-contrast background, one may naturally ask the following questions. Is there an advantage in describing these displays in terms of transparency? And, why should the perceptual construction of transparency predict a lowering in apparent contrast? We consider these questions in turn.

One advantage of describing the displays in Figure 20, a and c, in terms of perceptual transparency is that such a description automatically captures the surface layout and surface attributes (such as opacity) perceived in these displays. In addition, the transparency description also provides a computational rationale (in the sense of Marr, 1982) for the dependence of contrast–contrast on a number of spatial parameters to which it has been shown to be sensitive. Chubb et al. (1989) showed, for example, that the magnitude of the effect is greatly diminished when the texture patch and its surround differ in spatial frequency by more than one octave. Similarly, Solomon et al. (1993) showed that the induction is strongest when the gratings in the texture patch and its surround have the same orientation and weakest when they are perpendicular. More recently, Za, Iverson, and D'Zmura (1997) and Spehar, Arend, and Gilchrist (1995) have shown that the induction is greatest when the gratings in the texture patch and its surround have the same phase and weakest when they are completely out of phase (i.e., contrast polarity reversed). Low-level accounts of contrast–contrast explain these sensitivities by postulating that texture-sensitive units are tuned to specific spatial frequencies and orientations so that, for example, activity in a unit tuned to a certain frequency range most strongly affects the gain of

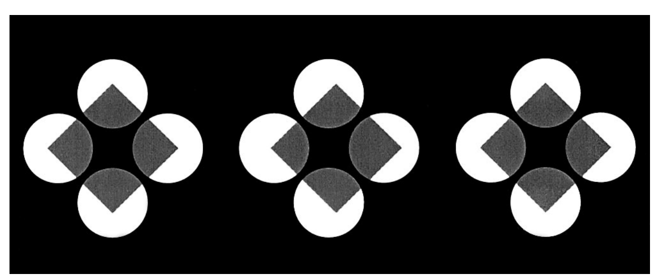

Figure 22. A demonstration of the effect of perceived transparency on brightness (display based on Nakayama et al., 1990; Varin, 1971). Depending on which two half-images are fused, this stereogram generates two different percepts: an opaque diamond seen through four portholes or a transparent diamond overlying four white disks. Note that the diamond appears darker when it is seen as a transparent surface than when it is seen as opaque. This is because, when it is seen as transparent, part of its lightness is attributed to the underlying white disks that are seen through it. From "Transparency: Relation to Depth, Subjective Contours, Luminance, and Neon Color Spreading," by K. Nakayama, S. Shimojo, & V. Ramachandran, 1990, *Perception, 19,* p. 502, Figure 6. Copyright 1990 by Pion Limited, London. Adapted with permission.

those neighboring units that are tuned to the same range, and only weakly to other units (Chubb et al., 1989; Solomon et al., 1993). Although these dependencies can thus be built into the neural models for the phenomenon, this does not in itself provide a computational rationale for these dependencies. Why should contrast–contrast be sensitive to spatial frequency, orientation, and phase?

To answer this question, consider the following quote by D'Zmura and Singer (1999), who have developed one of the most elaborate models of the lateral-inhibition account of contrast–contrast:

> Perhaps the greatest weakness of the model is that it provides no role for higher-level, perceptual mechanisms in contrast induction. The contrast-modulated White's illusions . . . show that these higher-level mechanisms alter apparent contrast, and one must wonder just how much of the processing that has been assigned to low-level mechanisms is best placed there. In viewing [figures displaying the selectivity of contrast–contrast to spatial frequency, orientation, phase, and chromatic properties], for instance, one notes that contrast induction is accompanied by the perception of transparency. *The central disks that have the weakest apparent contrast appear as though a Metelli-type filter has been placed atop a much larger disk. Are spatial frequency, orientation and color selectivities in fact properties of transparency-sensitive mechanisms that are able to separate an image into two layers for transparent overlay and underlying surfaces?* [italics added] (pp. 370–371)

The fact that displays that exhibit the strongest induction ("weakest apparent contrast") are also the ones that generate the percept of transparency is entirely consistent with the criterion of figural grouping we articulated for perceiving transparency—because making the center and surround differ in spatial frequency, orientation, or phase disrupts the textural continuity that is required for decomposing an image region into separate surfaces. Thus, as D'Zmura and Singer suggested, these parameters can be viewed as—and are naturally united within the framework of—low-level ingredients of figural grouping that are required to initiate percepts of transparency. From this perspective, the visual system contains texture-sensitive units that are tuned to these spatial parameters, at least in part, because these parameters play an important role in determining figural grouping that is necessary for computing surface structure.

The above argument is not intended to suggest that there are no low-level contributions to contrast–contrast. When the center and surround in a display of contrast–contrast are made to differ in their spatial frequency, orientation, or phase (so that these displays are no longer consistent with transparency) there is still some induced reduction in apparent contrast (see, e.g., Chubb et al., 1989; Solomon et al., 1993). Nevertheless, it remains true that displays that exhibit the strongest induction are invariably the ones that are consistent with transparency. In cases in which the display is not consistent with transparency, modifying it suitably so that it does become consistent with transparency (e.g., by aligning the gratings in the center and surround, preserving their contrast polarity, etc.) typically increases the size of the illusion. Computations of transparency may thus be at least partly responsible for the contrast–contrast illusion.

A final question remains: Why is the percept of transparency in a textured image region associated with a reduction in apparent contrast in that region? If the visual system were attempting to estimate the contrast of the underlying surface, for example, might we not expect an increase, rather than a decrease, in apparent contrast? The answer to this question is entirely analogous to related phenomena in the luminance domain. In particular, when the boundary of the test patch corresponds precisely to that of the perceived transparent surface, the brightness of the patch is biased toward the lightness of this transparent surface (recall Figure 22). In textured displays such as in Figure 20, a and c, the contrast border that defines the target patch is also the boundary of the perceived transparent surface. Hence, as in Figure 22, judgments of local brightness should be biased toward the lightness of the transparent layer. In other words, the light gray regions in the central patch should appear darker (relative to situations in which a scission into separate layers does not occur, e.g., Figure 20, b and d), because part of their lightness is attributed to the underlying white surface. Similarly, the dark gray regions should appear lighter in Figure 20, a and c, because part of their darkness is attributed to the underlying black surface. However, this is precisely equivalent to saying that the contrast in the central regions should be perceived as being lower in Figure 20, a and c, than in Figure 20, b and d, respectively. Hence the perceptual construction of transparency in a contrast-modulated display does, in fact, predict a lowering in apparent contrast in the region of transparency. Using this fact, we can now return to the second-order White's illusion (shown in Figure 21a) and see how the scission account also explains this effect. Indeed, one need only note that the four texture patches on top scission perceptually into two separate layers, a gray transparent layer and underlying high-texture stripes, whereas the patches below do not undergo such scission. Consistent with this, it is the patches on top that appear to have lower contrast—even though the reverse effect would be predicted by lateral inhibition.

The link between contrast induction and transparency can be demonstrated even in the case of standard Metelli-type displays. For example, compare the Metelli display in Figure 23a with a version that has a homogeneous surround (Figure 23b). The contrast in the central region of the transparency display is perceived to be lower. (This remains true also if, instead of making the surround homogeneous, we disrupt the figural conditions for transparency; compare, for example, the two displays in Figure 3, a and b.) The reduction in apparent contrast in Figure 23a relative to Figure 23b cannot be attributed simply to simultaneous contrast (viz., to the argument that the dark gray region in Figure 23a looks lighter because it is bounded by more black, and the light gray region looks darker because it is bounded by more white). One can greatly decrease the contribution of simultaneous contrast by shrinking the lengths of the contours over which simultaneous contrast can operate (see Figure 23c), but the contrast induction remains. Moreover, simultaneous contrast predicts that the dark gray bar at the extreme left of the central texture patch (in Figure 23c) should appear darker than the other dark gray bars (because it is bounded by more white) and that the light gray bar at the extreme right should appear lighter than the other light gray bars (because it is bounded by more black), but these predictions are not borne out perceptually.

In sum, we agree that the central regions in the displays on the left of Figure 20 appear to have lower contrast, and there clearly exist low-level mechanisms of contrast normalization. We have argued, however, that this description and explanation are incom-

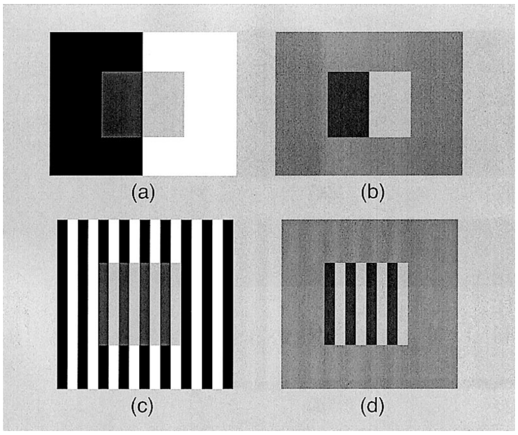

Figure 23. (a) Demonstrating the link between contrast–contrast and transparency with a standard Metelli display. The central region in (a) appears to have lower contrast than in (b), which has a homogeneous surround. According to the transparency-based account, this is because the dark gray region in (a) looks lighter (because part of its darkness is attributed to the underlying black) and the light gray region in (a) looks darker (because part of its lightness is attributed to the underlying white). The lowering of apparent contrast in (a) cannot be attributed simply to simultaneous contrast. In (c), the contours over which simultaneous contrast can operate have been significantly shrunk, but the contrast induction remains. Moreover, simultaneous contrast predicts that the left-most and right-most bars should look different than the other bars, because these two bars are more susceptible to effects of simultaneous contrast. However, this prediction is not borne out perceptually.

plete. The description of the phenomenology associated with standard contrast–contrast displays is incomplete because another salient perceptual difference between the display pairs in Figure 20 is the difference in perceived transparency. Similarly, the explanation based on lateral inhibition is incomplete because—as in the luminance domain—it does not take into account layered surface representations that are subsequently constructed at later stages of visual processing. As a result, in some notable cases (such as White's illusion; see Figure 21a), lateral inhibition actually predicts the incorrect sign of the illusion. We have argued that the perceptual construction of transparency is at least partly responsible for the contrast–contrast illusion (see also D'Zmura & Singer, 1999). Moreover, describing the perceptual difference between these displays in terms of transparency has the advantage of (a) capturing perceived surface structure and surface attributes (such as perceived opacity); (b) providing an ecological rationale for the sensitivity of contrast phenomena to spatial frequency, orientation, and phase; and (c) predicting the correct sign of the effect in White's illusion, where lateral inhibition fails.

Relationship to Previous Empirical Work on Transparency

Our results demonstrate that perceptual transparency deviates systematically from the predictions of Metelli's model—in terms of both when to initiate image scission to create the percept of transparency and how to assign perceived surface attributes to transparent layers. The results of Experiment 1 show that observers are able to make consistent and reliable matches of perceived transmittance (or opacity) of the transparent layer. These matches are determined by the Michelson contrast in the region of transparency relative to the contrast in adjoining regions. This result is not consistent with the solutions derived from Metelli's equations or their natural extensions—as these predict that perceived transmittance should be determined by relative luminance differences (or, more generally, luminance range). The results of Experiment 2 show that although observers are also able to judge the surface lightness or "color" of the transparent layer, this task is considerably more difficult. When observers attempt to make these matches, their responses are biased strongly toward the mean luminance in the region of transparency. As a result, these matches are consistent with only an incomplete discounting of the underlying layer. Finally, the results of Experiment 3 show that the "lowering of contrast" required for initiating the percept of transparency cannot be defined in terms of lowered luminance differences (or, more generally, luminance range), as predicted by Metelli's magnitude constraint. A clear failure of this constraint occurs in the regime in which luminance range is lowered, but Michelson contrast is increased, relative to adjoining regions. Observers do not perceive transparency in this regime despite the lowered luminance range. We have thus argued that variation in luminance contrast over continuous textures is a critical image property that the visual system uses to initiate the percept of transparency. Once such contrast variation is detected, the visual system scissions the lower contrast regions into multiple layers: an underlying surface that groups with the high-contrast background and an overlying transparent layer.

In this section, we review previous empirical work aimed at testing the predictions derived from the episcotister model, and we compare it to our own results. Beck et al. (1984) tested the qualitative constraints that follow from the restrictions on Metelli's solutions (Equations 3 and 4), namely, $0 < \alpha < 1$ and $0 < t < 1$. (Note that t is restricted to be less than 1 because it is taken to be reflectance value.) They began with four fixed reflectance values and permuted them in different ways to create stimulus configurations like those in Figure 1b and Figure 13. They then presented these to participants and asked them to judge whether or not they perceived transparency in these displays. On the basis of participants' responses, they concluded that violations of the two qualitative constraints based on the restriction $0 < \alpha < 1$ on the solution in Equation 3 adversely affect the perception of transparency—namely, Constraint 1, $p > q \Leftrightarrow a > b$; Constraint 2, $|p - q| < |a - b|$. However, violations of the qualitative constraints derived from the restriction $0 < t < 1$ on the solution in Equation 4 for t do not hinder the perception of transparency—namely, Constraint 3, $aq - bp > 0 \Leftrightarrow a + q - b - p > 0$; and Constraint 4, $|aq - bp| < |a + q - b - p|$.

On the basis of these results, Beck et al. (1984) made two proposals. First, they proposed alternative generative equations for transparency based on a "filter model," rather than the episcotister model. In this model, they included the effects of repetitive reflections between the transparent layer and the underlying surface. They argued that the first and second qualitative constraints (expressed in terms of reflectance values) are predictive of perceived transparency because they are consistent with the predictions of both the episcotister model and their filter model.

Second, they proposed that the terms a, b, p, and q in the solutions in Equations 3 and 4 for the perceived α and t should be taken to be lightness values, not reflectance values. In an experiment involving "partial transparency" (a situation in which, putatively, only the region of overlap in a stimulus configuration like that in Figure 13 is seen as transparent; see Metelli, 1974a), they found that participants' ratings of "index of transparency" are best predicted if lightness values are used in the solution (Equation 3) for α. They proposed that Constraints 1 and 2 might thus be evaluated by the visual system in terms of lightness values. Hence, they argued, because Constraints 1 and 2 involve ordinal comparisons of lightness values and lightness differences, these are easily checked by the visual system. On the other hand, Constraints 3 and 4 involve additions and multiplications of lightness values, and these are thus not easily interpretable.

Contrary to Beck et al.'s (1984) conclusion regarding the perceptual validity of Constraints 1 and 2, we found in Experiment 3 that Metelli's second qualitative constraint (the magnitude constraint) is not predictive of perceived transparency. In particular, observers consistently failed to see the central region in our displays as transparent when its luminance range was lower, but Michelson contrast higher, than the surround.[6] This was so despite the fact that the filter model and the episcotister model both predict the first and second constraints. Indeed, there is no difference, in principle, between generative equations derived from a physical setup involving an episcotister and that involving a transparent filter—other than the fact that the filter setup allows for a convenient way to model unbalanced transparency. The basic equations are the same in both cases (see Equations 5–7), involving both an attenuating multiplicative factor (the transmittance α) and an additive term. And in both cases, one may include repetitive reflections between the transparent layer and the underlying surface in a generative model (see Appendix A). One reason why Beck et al. may have considered the filter model to be fundamentally different is because the first term in their equations does not involve any dependence on α. However, this is due to a semantic difference in the usage of the term *reflectance:* By the *reflectance* of a screen, Beck et al. meant the total proportion of light reflected by a partially transmissive screen (effective reflectance f), whereas the usage consistent with Metelli's equations is the reflectance of just the material proportion of the screen (t; see Gerbino et al., 1990). As we pointed out in the Episcotister Model of Transparency section, these two usages are related by $f = (1 - \alpha)t$. With this substitution, Beck et al.'s equations look similar to those based on an episcotister—once repetitive reflections between the episcotister and the underlying surface are taken into account (see Equations A5 and A6 in Appendix A).

Beck et al.'s (1984) proposal that the terms a, b, p, and q should be taken to be lightness values is interesting. Its main advantage is that it captures the perceptual fact that if two light patches have the same reflectance difference as two dark patches, the light patches appear more similar (i.e., closer in brightness) than the dark ones. Indeed, it is possible that Metelli's second qualitative constraint expressed in terms of lightness values does predict when transparency is perceived. However, Beck et al. analyzed their yes–no transparency judgments (from their Experiment 1) only in terms of reflectance values and not in terms of lightness values. Such a proposal may also, in principle, account for the systematic deviations between perceived transmittance and the predictions derived from generative equations (which are based on luminance differences). However, this deviation is also captured by any compressive nonlinearity between image luminance and perceived brightness. Indeed, a disadvantage of treating a, b, p, and q as lightness values is that this view presupposes the existence of an initial stage where lightness values are computed locally in the image, prior to the subsequent construction of perceived surfaces (see Gerbino et al., 1990), whereas one typically thinks of lightness as an attribute assigned to perceived surfaces. Moreover, as Beck et al. argued, treating a, b, p, and q as lightness values makes any qualitative constraint based on the solution for t perceptually meaningless. However, as we showed in the When to Scission: I. Luminance Conditions for Transparency section, treating these terms as luminance values provides a natural way of interpreting the image constraint based on the solution for T. Finally, in a subsequent experiment involving "complete transparency," Beck et al. found that an α based on lightness values did no better in predicting participants' ratings of perceived transmittance than one based on reflectance values. They concluded that "We do not as yet have a good understanding of the factors controlling the judgment of transparency [transmittance] with complete transparency" (p. 420). Our own results indicate that an α based on the ratio of Michelson contrasts provides an excellent predictor of perceived surface transmittance.

Metelli et al. (1985) provided an alternative interpretation of Beck et al.'s (1984) result that violations of Qualitative Constraints 1 and 2 disrupt the perception of transparency, whereas violations of Constraints 3 and 4 do not. They argued that because these four constraints are derived from the episcotister model, they are valid only in the context of balanced transparency. (Recall that Solutions 3 and 4 for α and t can be derived only under the assumption that the transparent layer is balanced.) Under the assumption of balanced transparency, violations of Constraints 1 and 2 imply that α does not lie between 0 and 1, and violations of Constraints 3 and 4 imply that t does not lie between 0 and 1. Metelli et al. argued that, as both of these implications are physically meaningless, the assumption of balanced transparency must be false in these cases. In other words, displays that violate Conditions 1 and 2 are still consistent with transparency, but the transparency is unbalanced in transmittance (i.e., the transparent layer has different transmittances in regions **P** and **Q**). Similarly, displays that violate Conditions 3 and 4 are still consistent with transparency, but the transparency is unbalanced in reflectance (i.e., the transparent layer has different reflectances in regions **P** and **Q**). They concluded that violations of Constraints 1 and 2 disrupt the perception of transparency because human observers are not able to see transparency that is unbalanced in transmittance. However, violations of Constraints 3 and 4 do not disrupt the perception of transparency because human observers are able to see transparency that is unbalanced in reflectance.

The above conclusion is clearly false. Human observers readily see transparency that is unbalanced in transmittance. Indeed, this is the most common form of transparency found in nature (e.g., in scenes with fog, haze, smoke, or clouds). Part of the problem with the above arguments is that they are based solely on four-intensity

[6] Note that luminance range is a constant multiple of reflectance range under uniform illumination. Hence, Constraint 2 expressed in terms of luminance values is equivalent to Constraint 2 expressed in terms of reflectance values: $P - Q < A - B \Leftrightarrow pI - qI < aI - bI \Leftrightarrow p - q < a - b$.

displays. As we pointed out earlier, it is difficult to judge the lightness and transmittance of the transparent surface in such displays, because this surface is defined by just two regions of distinct luminance. The use of such displays is especially problematic for drawing conclusions about unbalanced transparency because surface qualities of the transparent layer now have to be assigned on the basis of a single image region with constant luminance (because the portions of the transparent surface corresponding to image regions **P** and **Q** can differ in transmittance and/or reflectance). Moreover, the interpretation of unbalanced transparency in these displays involves a highly nongeneric (or accidental) placement of the transparent surface and/or viewing position, because the locus of discontinuity in transmittance or reflectance (i.e., the **P**–**Q** border) is aligned precisely with the contour on the underlying surface. Using a richer class of displays, one sees that gradations in luminance contrast are easily interpreted by the visual system as inhomogeneities in the transmittance of the transparent layer (recall Figure 5).

Gerbino et al. (1990) used a matching experiment to study how human vision assigns lightness to transparent layers. On any given trial, all luminance values of the test display (A, B, P, Q) as well as the background luminances of the matching display (A', B') were fixed. The participants' task was to adjust the luminances P' and Q' in the center of the matching display so as to make the transparent layers in the two displays look identical. The values P' and Q' varied together so as to keep the value of α' (as determined by the episcotister model) fixed and equal to the value of α in the test display: $\alpha = (P - Q)/(A - B) = (P' - Q')/(A' - B') = \alpha'$. On the basis of the pattern of matches, Gerbino et al. concluded that the best predictor of perceived lightness of the transparent layer is provided by effective luminance $F = (1 - \alpha)T$, as given by the luminance version of the episcotister model. This is essentially the additive term in Metelli's equations: $P = \alpha A + (1 - \alpha)T = \alpha A + F$ (recall Equation 5).

As we noted earlier when discussing the results of Experiment 2, effective luminance fails to capture the qualitative trend in our lightness-matching data (its predictions are plotted in the bottom panel of Figure 12). Predictions made based on Metelli's term T, on the other hand, do capture this qualitative trend—although, quantitatively, they predict a stronger convergence than observers actually exhibit. Why do Gerbino et al.'s (1990) conclusions differ from our own? This may have to do in part with the fact that in their experiment, the values P' and Q' (to be adjusted by the participants) were made to covary so as to keep the transmittances equal in the test and matching displays, that is, $\alpha = \alpha'$. This was done to control for transmittance and study how lightness is assigned to transparent layers. (Their participants were instructed to make the two transparent layers look identical—not simply to match their lightnesses, while ignoring their transmittances.) However, Gerbino et al. assumed that perceived transmittance is given by the prediction of the episcotister model, namely, by the ratio $(P - Q)/(A - B)$; they did not explicitly test this assumption. As we saw in our Experiment 1, perceived transmittance is not determined by the predictions of Metelli's equations. It is consistent with the ratio of Michelson contrasts, rather than the ratio of luminance ranges or luminance differences. Moreover, although Gerbino et al. contrasted the predictions of effective luminance F against predictions derived from other hypotheses that do not invoke the notion of scission (e.g., matches based purely on image luminances in the region of transparency, or matches based on simultaneous contrast), they did not compare predictions based on F with those based on T. Hence the possibility remains open that T may in fact provide a better predictor of their own data than F.

An important contribution of Gerbino et al.'s (1990) article is that it convincingly argued that luminance, rather than reflectance or lightness, is the appropriate variable on which theories of transparency should be built. Furthermore, as we saw in the Episcotister Model of Transparency section, they showed that generative equations formally identical to Metelli's equations can be derived in the luminance domain. We agree that luminance is by far a better choice than reflectance or lightness, because these latter variables are surface attributes that must be inferred by the visual system (once visual surfaces have been constructed) based on the pattern of luminance values in the image. However, we want to stress the role of luminance *contrast* as the basic variable that the visual system tracks in order to compute transparency—both in order to initiate a perceptual decomposition into multiple surfaces in depth and to assign perceived transmittance to the transparent surface.

D'Zmura et al. (1997) extended Metelli's equations to three-dimensional color space (see also Da Pos, 1989; Faul, 1997; and Nakauchi, Silfsten, Parkkinen, & Ussui, 1999, for other work on color transparency). They showed that placing a transparent filter over a mosaic of colored patches leads to a convergence (or, in the limiting case, a translation) in color space, relative to the situation in which the same patches are seen in plain view. Consistent with this, they argued that coherent convergence/translation in color space along the boundary of an image region generates the percept of transparency in that region, whereas other kinds of transformations (such as rotations and shears) do not. Moreover, they showed that equiluminant convergence and translation in color space can also elicit an impression of transparency—even though these cannot be generated physically by the presence of an intervening episcotister or filter. This provides another reason for studying perceptual transparency as an independent phenomenon, rather than equating it with the inverse of physical models.

In a test of the convergence model, Chen and D'Zmura (1998) presented observers with Metelli-type configurations (recall Figure 1b) containing four different colors. Three of these colors were set in advance, and the fourth had to be set by observers in order to make the central region appear transparent. The convergence model predicts that the settings of the fourth color should lie on a line segment in color space determined by the other three colors. For the most part, observers' settings conformed to this prediction. However, the data also revealed two kinds of deviations. First, observers avoided settings that led the two halves of the central region to have complementary hues—even if these were predicted by the convergence model. This is consistent with an observation made earlier by Da Pos (1989) that the two regions that constitute the region of putative transparency must share a hue. Second, observers' settings deviated from the predictions of the convergence model in regions of color space that are perceptually nonuniform (i.e., where hue changes rapidly in some subregions and slowly in others). As in the luminance domain, Metelli's equations applied to the color domain do not take into account this perceptual inhomogeneity. On the basis of our own results, we suggest that a natural interpretation of this deviation is that perceived color contrast, rather than absolute differences in physically defined color space, is the critical variable that the visual system uses to compute perceptual transparency. Moreover, as the results of our

Experiment 3 indicate, some convergences, namely, those that do not lead to a lowering in Michelson contrast (recall the regime $R-C+$ in Figure 15b) do not generate percepts of transparency.

More recently, D'Zmura, Rinner, and Gegenfurtner (2000) conducted a matching experiment in which observers adjusted the color of surface patches seen behind a transparent filter to match the color of corresponding patches seen in plain view. They found that the convergence model with only 4 parameters (a scalar for convergence and a three-dimensional vector for translation in color space) fit the data almost as well as the full affine model with 12 parameters. On the basis of the fits of the convergence model, they inferred that observers underestimate the transmittance of the filter by a factor of almost two (see also Hagedorn & D'Zmura, 2000). Our own results (Experiment 1) indicate that the relationship between physical transmittance and perceptual transmittance is more complex: Observers systematically overestimate the transmittance of darkening transparent surfaces and systematically underestimate the transmittance of lightening transparent surfaces. Moreover, the magnitude of this deviation (i.e., overestimation or underestimation) is directly proportional to the extent to which the transparent layer changes the mean luminance projected from the underlying surface. However, we agree entirely with D'Zmura et al.'s (2000) overall conclusion that observers can discount only partially the color shifts introduced by the presence of an intervening transparent layer. Although D'Zmura et al.'s (2000) experiment is very different from our Experiment 2 (in our experiment, observers matched the lightness of the transparent layer itself, not of the underlying surface), both experiments point to the same conclusion, namely, that observers can separate only incompletely the contributions of an opaque surface and an overlying transparent layer.

Conclusions

Theories of perceptual transparency have typically been tied to physical (or generative) models that lead to percepts of transparency. However, the relationship between perceptual theory and generative model is often left rather vague. Usually, it is implicitly assumed that perceptual transparency is determined by inverting the generative equations. Our results indicate that perceptual transparency deviates systematically from the predictions of Metelli's equations and their natural extensions—in terms of both when to scission an image region into multiple surfaces and how to assign attributes to transparent surfaces. This underscores the need to distinguish clearly between generative models of transparency and perceptual theories, so that the relationship between them may be clearly articulated.

The results of Experiment 1 demonstrate that Michelson contrast is the critical image variable that the visual system uses to assign perceived transmittance to transparent surfaces. This result is not consistent with the solutions derived from Metelli's equations—which predict that perceived transmittance should be determined by relative luminance differences (or, more generally, luminance ranges). The usage of Michelson contrast leads human observers to systematically overestimate the transmittance of transparent layers that darken the underlying surfaces and to systematically underestimate the transmittance of transparent layers that lighten the underlying surfaces. Thus, the perceived transmittance of a transparent surface depends not only on its own physical transmittance but also on the mean luminance projected from the underlying surfaces that it partially occludes.[7]

Ironically, whereas most studies of perceptual transparency have argued for the perceptual validity of Metelli's equations (Beck et al., 1984, is a notable exception), Metelli himself realized that there was something wrong with the predictions of his equations. Metelli (1974a) noted an observation that a black episcotister looks more transmissive than a white episcotister of the same physical transmittance α—whereas his model would predict both episcotisters to look equally transmissive. However, he never reconciled this observation with his formal model. Our proposal that the expression for transmittance α should be rewritten as a ratio of Michelson contrasts resolves this long-standing puzzle, because the black episcotister generates a higher contrast than the light one.

The results of Experiment 2 demonstrate that in estimating the lightness of a transparent surface, human observers are able to discount only partially the luminance generated by the underlying surfaces. As a result, lightness matches are biased strongly toward the mean luminance in the region of transparency—again deviating from the predictions based on Metelli's equations. Finally, the results of Experiment 3 demonstrate that, contrary to Metelli's magnitude constraint (recall Condition 2), the visual system does not initiate percepts of transparency based on lowered luminance range. In particular, observers do not perceive transparency in regimes in which the luminance range is lowered, but Michelson contrast is increased, relative to adjoining image regions. We thus argued that the visual system uses changes in Michelson contrast over aligned contours and groupable textures as a critical image property to initiate percepts of transparency. Once such contrast variation is detected, the visual system scissions the lower contrast regions into multiple surfaces: an underlying surface that groups with the high-contrast background and an overlying transparent surface.

[7] We have demonstrated in recent work that, in the case of translucent surfaces that scatter light as well (a situation that Metelli's equations do not model; recall Figure 19c), the degree of blur provides a further image cue—above and beyond Michelson contrast and apparent contrast—that the visual system uses to assign surface opacity (Singh & Anderson, 2002).

References

Adelson, E. H. (1993, December 24). Perceptual organization and the judgment of brightness. *Science, 262,* 2042–2044.

Adelson, E. H. (2000). Lightness perception and lightness illusions. In M. Gazzaniga (Ed.), *The new cognitive neurosciences* (pp. 339–351). Cambridge, MA: MIT Press.

Adelson, E. H., & Anandan, P. (1990, July). *Ordinal characteristics of transparency.* Paper presented at the AAAI–90 Workshop on Qualitative Vision, Boston, MA.

Albert, M. K., & Hoffman, D. D. (1995). Genericity in spatial vision. In R. D. Luce, M. D'Zmura, D. D. Hoffman, G. J. Iverson, & A. K. Romney (Eds.), *Geometric representations of perceptual phenomena* (pp. 95–112). Mahwah, NJ: Erlbaum.

Anderson, B. L. (1997). A theory of illusory lightness and transparency in monocular and binocular images: The role of junctions. *Perception, 26,* 419–453.

Anderson, B. L. (1999). Stereoscopic surface perception. *Neuron, 24,* 919–928.

Beck, J., & Ivry, R. (1988). On the role of figural organization in perceptual transparency. *Perception & Psychophysics, 44,* 585–594.

Beck, J., Prazdny, K., & Ivry, R. (1984). The perception of transparency with achromatic colors. *Perception & Psychophysics, 35,* 407–422.

Biederman, I. (1987). Recognition-by-components: A theory of human image understanding. *Psychological Review, 94,* 115–147.

Blinn, J. F. (1982). Light reflection functions for simulation of clouds and dusty surfaces. *Computer Graphics, 16,* 21–29.

Chen, V., & D'Zmura, M. (1998). Test of a convergence model for color transparency perception. *Perception, 27,* 595–608.

Chubb, C., Sperling, G., & Solomon, J. (1989). Texture interactions determine perceived contrast. *Proceedings of the National Academy of Science USA, 86,* 9631–9635.

Da Pos, O. (1989). *Trasparenze* [Transparency]. Padua, Italy: Icone.

D'Zmura, M., Colantoni, P., Knoblauch, K., & Laget, B. (1997). Color transparency. *Perception, 26,* 471–492.

D'Zmura, M., Rinner, O., & Gegenfurtner, K. R. (2000). The colors seen behind transparent filters. *Perception, 29,* 911–926.

D'Zmura, M., & Singer, B. (1999). Contrast gain control. In L. T. Sharpe & K. R. Gegenfurtner (Eds.), *Color vision: From genes to perception* (pp. 369–385). Cambridge, England: Cambridge University Press.

Faul, F. (1997). *Theoretische and experimentelle Untersuchung chromatischer Determinanten perzeptueller Transparenz* [Theoretical and experimental study of the chromatic determinants of perceptual transparency]. Unpublished doctoral dissertation, Christian-Albrechts-Universitat zu Kiel, Kiel, Germany.

Freeman, W. T. (1994). The generic viewpoint assumption in a framework for visual perception. *Nature, 368,* 542–545.

Gerbino, W. (1994). Achromatic transparency. In A. L. Gilchrist (Ed.), *Lightness, brightness, and transparency* (pp. 215–255). Hove, England: Erlbaum.

Gerbino, W., Stultiens, C., Troost, J., & de Weert, C. (1990). Transparent layer constancy. *Journal of Experimental Psychology: Human Perception and Performance, 16,* 3–20.

Gilchrist, A. L. (1977, January 14). Perceived lightness depends on perceived spatial arrangement. *Science, 195,* 185–187.

Hagedorn, J., & D'Zmura, M. (2000). Color appearances of surfaces viewed through fog. *Perception, 29,* 1169–1184.

Hanrahan, P., & Krueger, W. (1993). Reflection from layered surfaces due to subsurface scattering. *Proceedings of the 20th annual Conference on Computer Graphics and Interactive Techniques* (pp. 165–174). New York: ACM Press.

Helmholtz, H. V. (1962). *Helmholt's treatise on physiological optics.* New York: Dover. (Original work published 1866)

Hochberg, J. E., & Beck, J. (1954). Apparent spatial arrangement and perceived brightness. *Journal of Experimental Psychology, 47,* 263–266.

Kanizsa, G. (1979). *Organization in vision.* New York: Praeger.

Kasrai, R., & Kingdom, F. A. A. (2001). Precision, accuracy, and range of perceived achromatic transparency. *Journal of the Optical Society of America, A, 18,* 1–11.

Knill, D., & Kersten, D. (1991). Apparent surface curvature affects lightness perception. *Nature, 351,* 228–230.

Koenderink, J. J., & van Doorn, A. J. (1979). The internal representation of solid shape with respect to vision. *Biological Cybernetics, 32,* 211–216.

Koffka, K. (1935). *Principles of Gestalt psychology.* New York: Harcourt, Brace.

Kubelka, P. (1954). New contributions to the optics of intensely light-scattering materials: II. Non-homogeneous layers. *Journal of the Optical Society of America, 44,* 330–335.

Loula, F., Kourtzi, Z., & Shiffrar, M. (2000). Surface segmentation cues influence implicit visual memory for novel and familiar shapes. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26,* 929–944.

Lowe, D. G. (1985). *Perceptual organization and visual recognition.* Boston: Kluwer Academic.

Marr, D. (1982). *Vision.* New York: Freeman.

Metelli, F. (1970). An algebraic development of the theory of perceptual transparency. *Ergonomics, 13,* 59–66.

Metelli, F. (1974a). Achromatic color conditions in the perception of transparency. In R. B. MacLeod & H. L. Pick (Eds.), *Perception: Essays in honor of J. J. Gibson* (pp. 95–116). Ithaca, NY: Cornell University Press.

Metelli, F. (1974b, April). The perception of transparency. *Scientific American, 230,* 90–98.

Metelli, F. (1985). Stimulation and perception of transparency. *Psychological Research, 47,* 185–202.

Metelli, F., Da Pos, O., & Cavedon, A. (1985). Balanced and unbalanced, complete and partial transparency. *Perception & Psychophysics, 38,* 354–366.

Nakauchi, S., Silfsten, P., Parkkinen, J., & Ussui, S. (1999). Computational theory of color transparency: Recovery of spectral properties for overlapping surfaces. *Journal of the Optical Society of America, A, 16,* 2612–2624.

Nakayama, K., & Shimojo, S. (1992, September 4). Experiencing and perceiving visual surfaces. *Science, 257,* 1357–1363.

Nakayama, K., Shimojo, S., & Ramachandran, V. (1990). Transparency: Relation to depth, subjective contours, luminance, and neon color spreading. *Perception, 19,* 497–513.

Richards, W., & Stevens, K. (1979, December). *Efficient computations and representations of visible surfaces* (Final Report, Contract No. AFOSR-79-0200). Unpublished report, Massachusetts Institute of Technology.

Singh, M., & Anderson, B. L. (2002). Perceptual assignment of opacity to translucent surfaces: The role of image blur. *Perception, 31,* 531–552.

Singh, M., & Hoffman, D. D. (1998). Part boundaries alter the perception of transparency. *Psychological Science, 9,* 370–378.

Solomon, J., Sperling, G., & Chubb, C. (1993). The lateral inhibition of perceived contrast is indifferent to on-center/off-center segregation, but specific to orientation. *Vision Research, 33,* 2671–2683.

Somers, D. C., & Adelson, E. H. (1997). Junctions, transparency, and brightness. *Investigative Ophthalmology and Vision Science, 38,* 2952.

Spehar, B., Arend, L. E., & Gilchrist, A. L. (1995). Contrast contrast: Interactions between spatial and luminance factors. *Review of Psychology, 2,* 3–12.

Spehar, B., & Zaidi, Q. (1997). New configurational effects on perceived contrast and brightness: Second-order White's effects. *Perception, 26,* 409–417.

Tipper, S. P. (1985). The negative priming effect: Inhibitory priming by ignored objects. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 37(A),* 571–590.

Todorovic, D. (1997). Lightness and junctions. *Perception, 26,* 379–394.

Tudor-Hart, B. (1928). Studies in transparency, form, and colour. *Psychologische Forschung, 10,* 255–298.

Varin, D. (1971). Fenomini di contrasto e diffusione cromatica nell' organizzazione spaziale del campo perettivo [Phenomena of contrast and chromatic diffusion in the spatial organization of the perceptual field]. *Rivista di Psicologia, 65,* 101–128.

Watanabe, T., & Cavanagh, P. (1993a). Surface decomposition accompanying the perception of transparency. *Spatial Vision, 7,* 95–111.

Watanabe, T., & Cavanagh, P. (1993b). Transparent surfaces defined by implicit X junctions. *Vision Research, 33,* 2339–2346.

White, M. (1979). A new effect of pattern on perceived brightness. *Perception, 8,* 413–416.

Witkin, A. P., & Tenenbaum, J. M. (1983). On the role of structure in vision. In J. Beck, B. Hope, & A. Rosenfeld (Eds.), *Human and machine vision* (pp. 481–543). Academic Press.

Za, S., Iverson, G., & D'Zmura, M. (1997). Phase selectivity in contrast induction. *Investigative Ophthalmology and Visual Science, 38,* 2952.

Zaidi, Q., Spehar, B., & Shy, M. (1997). Induced effects of backgrounds and foregrounds in three-dimensional configurations: The role of T-junctions. *Perception, 26,* 395–408.

Appendix A

Modification of Equations for the Case of Unequal Illumination of Layers

If the mesh or filter is either very large or very close to the background surface, the light source is able to illuminate the background only through the screen—so one can no longer assume that the background and screen are equally illuminated (the effective illumination of the background is now αI rather than I, as in Figure 4). By the time the light reflected by the background reaches the eyes, it has been attenuated by the screen twice. As a result, Equation 5 for the total amount of light reaching the eyes from region **P** is now modified as follows:

$$P = \alpha^2 aI + (1 - \alpha)tI = \alpha^2 A + (1 - \alpha)T. \quad \text{(A1)}$$

Similarly, the light reaching the eyes from region **Q** is given by the following:

$$Q = \alpha^2 B + (1 - \alpha)T. \quad \text{(A2)}$$

The solutions for α and t from this generative model are as follows:

$$\alpha = \sqrt{\frac{P - Q}{A - B}}, \quad \text{(A3)}$$

$$T = \frac{AQ - BP}{\sqrt{A - B}(\sqrt{A - B} - \sqrt{P - Q})}. \quad \text{(A4)}$$

In case the mesh or filter is very close to the background surface, interreflections between the two surfaces may also become significant. The amount of light P that reaches the eyes from region **P** is then given by the following:

$$\begin{aligned} P &= (1 - \alpha)tI + \alpha^2 aI + \alpha^2 a^2(1 - \alpha)tI + \alpha^2 a^3(1 - \alpha)^2 t^2 I + \cdots \\ &= (1 - \alpha)tI + \alpha^2 aI[1 + (1 - \alpha)at + (1 - \alpha)^2 a^2 t^2 + \cdots] \\ &= (1 - \alpha)tI + \alpha^2 aI/[1 - (1 - \alpha)at] \\ &= (1 - \alpha)T + \alpha^2 A/[1 - (1 - \alpha)at] \\ &= (1 - \alpha)T + \alpha^2 AI^2/[I^2 - (1 - \alpha)AT]. \quad \text{(A5)} \end{aligned}$$

Similarly, the amount of light Q that reaches the eyes from region **Q** is given by the following:

$$Q = (1 - \alpha)T + \alpha^2 B\left(\frac{I^2}{I^2 - (1 - \alpha)BT}\right). \quad \text{(A6)}$$

These equations differ from Beck et al.'s (1984) because (a) they wrote these equations purely in terms of reflectance values and (b) they used the term *reflectance* to mean the *effective reflectance* of a transparent layer, that is, the product $(1 - \alpha)t$, rather than its *material reflectance t* (i.e., the reflectance of the material portion of a mesh; see Gerbino et al., 1990).

Note that as the illuminant I gets large, the ratio $I^2/[I^2 - (1 - \alpha)BT]$ approaches unity, so that Equations A5 and A6 become identical to Equations A1 and A2, and the solutions for α and T therefore approach the solutions in Equations A3 and A4 (see also Gerbino, 1994).

Appendix B

Solutions for α and T for Textured Images

Consider a local patch over which $\alpha(x, y)$ and $T(x, y)$ are constant, with respective values α and T. Over this local patch, the mapping from $A(x, y)$ to $L(x, y)$ is given by the following:

$$L(x, y) = \alpha A(x, y) + (1 - \alpha)T, \quad \text{(B1)}$$

that is, a linear function with slope α and intercept $(1 - \alpha)T$. Because $L(x, y)$ is a monotonically increasing function of $A(x, y)$ ($\alpha > 0$), the minimum value of $A(x, y)$ maps on to the minimum value of $L(x, y)$, and the maximum value of $A(x, y)$ maps on to the maximum value of $L(x, y)$. Thus, over this local patch we have

$$L_{\min} = \alpha A_{\min} + (1 - \alpha)T, \quad \text{(B2)}$$

$$L_{\max} = \alpha A_{\max} + (1 - \alpha)T. \quad \text{(B3)}$$

Equations B2 and B3 can now be solved yielding solutions for α and T over this local patch:

$$\alpha = \frac{L_{\max} - L_{\min}}{A_{\max} - A_{\min}}, \quad \text{(B4)}$$

$$T = \frac{A_{\max} L_{\min} - A_{\min} L_{\max}}{(A_{\max} - A_{\min}) - (L_{\max} - L_{\min})}. \quad \text{(B5)}$$

Under the assumption of symmetrical distributions, that is,

$$L_{\text{mean}} = \frac{L_{\max} + L_{\min}}{2} \text{ and } A_{\text{mean}} = \frac{A_{\max} + A_{\min}}{2},$$

the expression for T in Equation B5 can be rewritten as follows:

$$T = \frac{A_{\text{range}} L_{\text{mean}} - A_{\text{mean}} L_{\text{range}}}{A_{\text{range}} - L_{\text{range}}}. \quad \text{(B6)}$$

(*Appendixes continue*)

Appendix C

Predictions of Luminance Difference and Michelson Contrast: Proofs

Result 1

For polarity preserving T junctions, luminance difference and Michelson contrast make the same predictions of lower contrast.

Proof

Polarity preserving T junctions can only have one of two forms: $A < P < Q$ or $P < Q < A$, where A is the luminance of the "top" of the junction (see Figure C1).

Case A: $A < P < Q$

It is clear that $P - A < Q - A$, so that P has a lower luminance difference with A than does Q. We must now show that the Michelson contrast between P and A is also lower than between Q and A; in other words,

$$\frac{P - A}{P + A} < \frac{Q - A}{Q + A}.$$

Consider the function

$$f(x) = \frac{x - A}{x + A}\,(A > 0).$$

Because

$$f'(x) = \frac{2A}{(x + A)^2} > 0,$$

$f(x)$ is an increasing function of x. Therefore,

$$P < Q \Rightarrow \frac{P - A}{P + A} < \frac{Q - A}{Q + A}.$$

Case B: $P < Q < A$

It is clear that $A - Q < A - P$. We must show that

$$\frac{A - Q}{A + Q} < \frac{A - P}{A + P}.$$

However, this follows immediately because

$$g(x) = \frac{A - x}{A + x} = -f(x)$$

is a decreasing function of x (for $A > 0$), and $P < Q$.

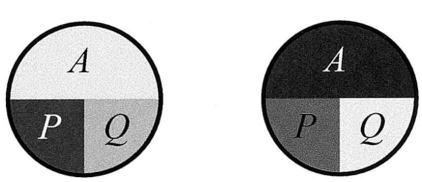

Figure C1. The two forms of polarity-preserving T junctions: The "top" of the junction (with luminance A) is either lighter than both of the other two regions or it is darker than both. For such T junctions, luminance difference and Michelson contrast make the same prediction of which half (left or right) has "lower contrast." This result does not hold more generally, that is, if the T junction is not polarity preserving.

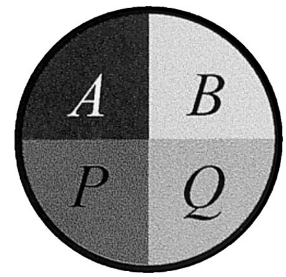

Figure C2. A single-reversing X junction: Only one of the two contours of the X junction (the vertical one here) preserves contrast polarity. For single-reversing junctions, luminance difference and Michelson contrast make the same prediction of which half (top or bottom) has "lower contrast." This result does not hold more generally, that is, if the X junction is either nonreversing or double reversing.

Note

Result 1 does not hold for T junctions that do not preserve contrast polarity. For a counterexample, consider a T junction defined by $A = 75$ cd/m^2, $P = 95$ cd/m^2, and $Q = 57$ cd/m^2. (As before, A refers to the luminance of the top of the junction.) Then the border **Q–A** has a lower luminance difference ($|A - Q| < |A - P|$), but the border **P–A** has a lower Michelson contrast:

$$\left(\frac{|A - P|}{|A + P|} < \frac{|A - Q|}{|A + Q|}\right).$$

Result 2

For single-reversing X junctions, luminance difference and Michelson contrast make the same predictions of lower contrast, across the edge that preserves polarity.

Proof

Single-reversing X-junctions are defined by four luminance values $A < P < Q < B$, where the regions with luminances A and Q are diagonally opposite from each other (see Figure C2). The contours defined by **A–B** and **P–Q** preserve contrast polarity, and the **P–Q** border has a lower luminance difference than the **A–B** border. We must show that **P–Q** border also has a lower Michelson contrast.

Because $P < Q < B$, it follows from Result I, Case A, that

$$\frac{Q - P}{Q + P} < \frac{B - P}{B + P}.$$

Similarly, because $A < P < B$, it follows from Case B that

$$\frac{B - P}{B + P} < \frac{B - A}{B + A}.$$

The above two inequalities now imply that

$$\frac{Q - P}{Q + P} < \frac{B - A}{B + A}.$$

Appendix D

Transparency Lowers Michelson Contrast: Proof for Textured Images

Over a locally balanced patch, the solution in Equation 11 for T can be rewritten as follows:

$$
\begin{aligned}
T &= \frac{A_{\text{range}}\,L_{\text{mean}} - A_{\text{mean}}\,L_{\text{range}}}{A_{\text{range}} - L_{\text{range}}} \\
&= A_{\text{mean}}\,L_{\text{mean}}\,\frac{\left(\dfrac{A_{\text{range}}}{A_{\text{mean}}} - \dfrac{L_{\text{range}}}{L_{\text{mean}}}\right)}{A_{\text{range}} - L_{\text{range}}} \\
&= 2A_{\text{mean}}\,L_{\text{mean}}\,\frac{A_{\text{contrast}} - L_{\text{contrast}}}{A_{\text{range}} - L_{\text{range}}}.
\end{aligned}
$$

(As before, A and L refer to luminances in the unobscured region and the region of transparency, respectively.) The restriction $\alpha \leq 1$ on the solution in Equation 10 already forces the denominator $A_{\text{range}} - L_{\text{range}}$ to be nonnegative. The restriction $T \leq 0$ thus forces the numerator $A_{\text{contrast}} - L_{\text{contrast}}$ to be nonnegative as well. In other words, both luminance range and Michelson contrast are necessarily lowered in the region of transparency.

Received January 19, 2000
Revision received June 27, 2001
Accepted June 27, 2001 ■