Precision, accuracy, and range of perceived achromatic transparency

Reza Kasrai

McConnell Brain Imaging Centre, Montréal Neurological Institute, 3801 rue University, Montréal, Québec H3A 2B4, Canada

Frederick A. A. Kingdom

McGill Vision Research Unit, 687 Avenue des Pins ouest, Room H4-14, Montréal, Québec H3A 1A1, Canada

Received February 11, 2000; revised manuscript received July 3, 2000; accepted July 14, 2000

How accurately do human observers perceive the properties of an achromatic transparent filter with both reflective and transmissive components? To address this question, a novel six-luminance stimulus was employed, consisting of three transparent layer luminances set against three background luminances, which satisfied the conventional constraints of perceptual transparency. In one experiment, subjects adjusted one of the three layer luminances to complete the impression of a uniform transparent disk. It was found that the luminance-based formulation of Metelli's episcotister model and a model based on ratios of Michelson contrasts best predicted the subjects' settings, which were both accurate and precise. In another experiment, pairs of stimuli selected from a range with various values of the adjustable layer luminance were presented in a series of forced-choice trials. A modified implementation of the pair comparisons method was employed to recover the distribution that describes each subject's preference pattern. Results showed that there exists a reasonably wide range of stimuli that give rise to at least some degree of perceived transparency.

OCIS codes: 330.7310, 330.4300, 330.5510, 330.4060.

1. INTRODUCTION

A. Reflectance Formulation

The perception of transparency has intrigued vision researchers for at least a hundred years. Investigators from Helmholtz[1] to Metelli[2] have proposed different theories to explain the phenomenal experience associated with perceiving a background through a transparent surface and employed different techniques for testing those theories. Metelli used an episcotister—a rotating opaque disk with open sectors of varying size giving the effect of different illusory opacities. When the alternation of the episcotister's open and opaque sectors exceeds the critical flicker frequency, it appears as a transparent disk through which a bipartite background can be seen, as in Fig. 1(b). Metelli's observations led to the formulation of his so-called episcotister reflectance model,[2,3]

$$p = ta + (1 - t)r, \qquad q = tb + (1 - t)r, \tag{1}$$

where r is the reflectance of the opaque episcotister surface, t is the fraction of the disk that is open, a and b are the reflectances of the background surfaces, and p and q are the resulting layer reflectances,[4] such that $0 \leqslant \{a, b, p, q, r, t\} \leqslant 1$. The episcotister reflectance model is an expansion of Talbot's law of color mixture, $\alpha x + (1 - \alpha)y = z$, where x and y are two colors expressed as reflectances, α and $1 - \alpha$ are the relative linear mixing proportions, and z is the resulting fusion color. In each case, reflectances from two sources contribute linearly to produce the desired mixture.

B. Illumination Considerations and Luminance Formulation

Metelli's original reflectance equations have since been reformulated by Gerbino[5] to allow for nonhomogeneous illumination of the transparent layer and the background, partly because illumination considerations cannot be incorporated into a purely reflectance-based formulation. It has been shown that the visual system is able to correctly perceive characteristics of surfaces in two depth planes that are differently illuminated.[6] Similarly, an opaque surface and a transparent overlay may in general be differently illuminated. This consideration necessarily requires reformulating Eqs. (1) in terms of luminances rather than reflectances. The luminance formulation of any transparency model is also more amenable to psychophysical studies by using CRT displays, where luminance (emitted pixel intensity) is the variable being manipulated. When luminance, instead of reflectance or lightness (perceived reflectance), is used as the independent variable to describe the necessary color relationships for perceptual transparency, Eqs. (1) become

$$P = taI + (1 - t)rI' = tA + (1 - t)rhI = tA + F$$

$$Q = tB + F, \tag{2}$$

where P and Q are the luminances of the portion of the image representing the layer (**P** and **Q**; see Fig. 1), A and B are the luminances of the background (**A** and **B**), I and I' are the (generally nonequal) illumination components for the two surfaces with ratio h, such that $I' = hI$, and t and r are the transmittance and the reflectance of the

0740-3232/2001/010001-11$15.00

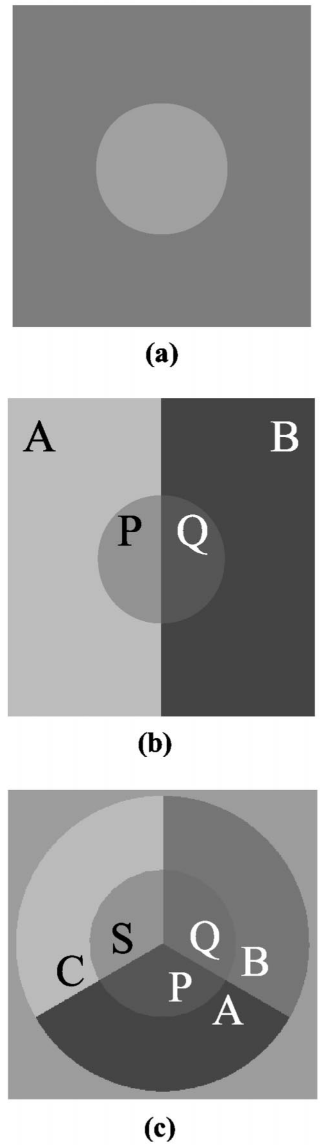

Fig. 1. Transparency stimuli: (a) two-luminance stimulus (does not provide a strong transparent percept), (b) four luminances simulating an episcotister in the classical bipartite display, (c) six-luminance stimulus with mid-gray background.

layer. The second term is collapsed into an overall additive component F, since with four known luminances (A, B, P, Q) and two equations, only two unknowns (t, F) can be extracted, and the product rhI cannot be disambiguated without additional assumptions or prior knowledge.

Figure 2 shows the optic array describing the luminance episcotister (LE) model from which Eqs. (2) are derived, which assumes that the background surface is illuminated directly and not through the layer. Others[7] have assumed that the background is illuminated exclusively through the transparent layer. From an ecological standpoint, isotropic diffuse illumination is the norm, and neither assumption is strictly valid. Rather, when the layer is close to or flush with the background, most of the light reflecting off the background (through the layer) will have already passed through the layer once, giving rise to a t^2 factor. This quadratic term, and multiple reflections between the filter and the background, are the defining characteristic of the filter model of Beck *et al.*[7] [the formulation for which is given in Subsection 3.B; see Eqs. (4)]. When the layer is at a distance from the background, most of the light reflected from the background will not have passed through the layer first, and Eqs. (2) are more valid. It should be made clear that even though these models have a basis in optical physics, they should be taken not as physical models *per se* but rather as models of the assumptions that the visual system makes when interpreting transparency stimuli.

Isolating t and F from the luminance-based Eqs. (2) gives

$$t = \frac{P - Q}{A - B}, \quad F = \frac{AQ - BP}{A - B}, \tag{3}$$

implying that given the four known luminances, the visual system can extract the values of t and F. Even though it has been shown that some form of illusory transparency is possible with just three luminances under certain conditions,[8,9] this is generally not the case; four luminances are necessary to unambiguously render a given combination of t and F, as evident from Eqs. (3). Figure 1 shows the progression from two luminances (no phenomenal transparency), to four luminances (the minimum necessary), to the six-luminance stimulus used in this study, the rationale for which is described in the next subsection.

C. Measuring Perceived Transparency

It is unlikely that the visual system calculates the values of t and F explicitly, as one does not appear to have phenomenal access to the absolute values of these parameters when viewing a transparency stimulus. This is not to say that one cannot perceive a filter to be, say, more opaque or less reflective than another, but it would presumably be difficult to select a filter with, say, $t = 0.5$ from an array of filters with different t values. How might one measure the accuracy and the precision with which human observers perceive and encode the properties of a simulated transparent surface? Accuracy is a measure of closeness to an expected setting. This may be a setting according to the observer's internal model of transparency. Precision, on the other hand, is the reliability or the reproducibility of the settings, which will be reflected in their variability. Consider the four-luminance display shown in Fig. 1(b). Fixing both of the background luminances and one of the layer luminances leaves a range of the other layer's luminances that will produce valid transparencies, i.e., valid combinations of t

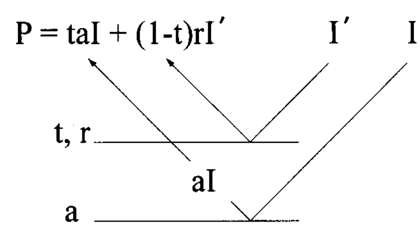

Fig. 2. Optic array describing the episcotister luminance model. Note that the background is illuminated only directly and not through the layer.

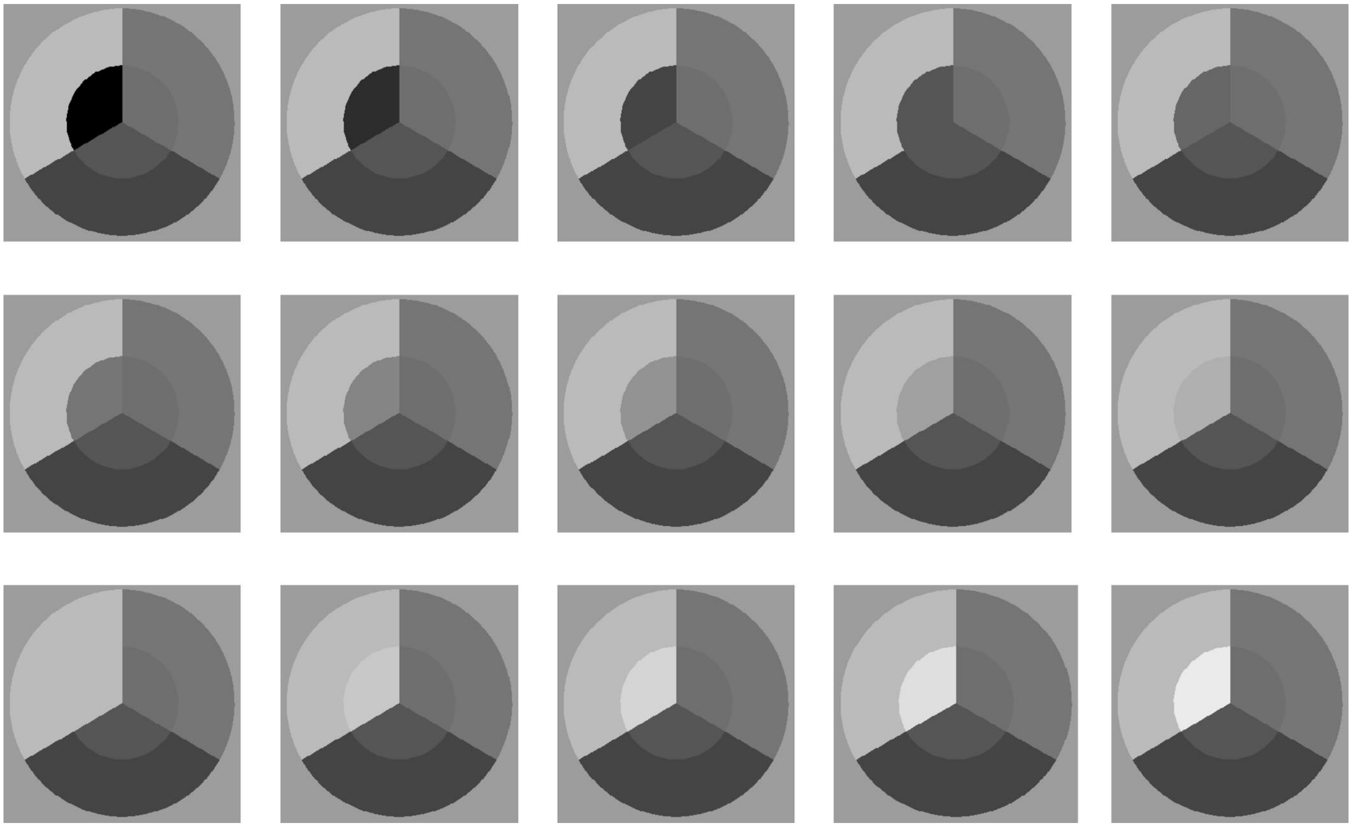

Fig. 3. Range of possible test patch settings. The correct image, according to the luminance episcotister (LE) model, is in the center.

and F. In other words, there is not a single solution with the four-luminance display, and therefore the four-luminance display is insufficient on its own for use as a tool for measuring perceived transparency. In their study of perceived transparency, Gerbino *et al.*[10] used *two* four-luminance displays (i.e., eight luminances in total): one a test stimulus, the other a comparison stimulus. Their subjects adjusted F (thus simultaneously altering both P and Q) on the test to perceptually match that of the comparison. The investigators showed that subjects were capable of making reasonably accurate settings according to the luminance version of the episcotister model.

A simpler and better suited stimulus for measuring transparency, however, is a six-luminance display,[11] one form of which is shown in Fig. 1(c). Six luminances are the minimum necessary for a unique solution with the use of just one adjustable luminance.[12] The stimulus in Fig. 1(c) consists of three background and three layer luminances. With two of the background luminances and their two corresponding layer luminances appropriately set, a valid transparent surface can be created (as the minimum four-luminance requirement for transparency is satisfied). With the third background luminance then set arbitrarily as a test, the subject can adjust the remaining layer luminance to complete the transparent surface—to create perceptually uniform t and F consistent with all three layer luminances. In principle, subjects do not need to be phenomenologically aware of the distinction between t and F with this method, since they are required only to make the whole layer appear uniform in its transparency properties. It is worth emphasizing here that subjects do not adjust t and F independently with this method. Consequently, the results do not directly show how well these two properties of transparent filters are independently set. Rather, the results speak to the overall salience of the transparency percept and how well subjects perform when manipulating a single stimulus luminance.

To understand how the six-luminance stimulus is employed, consider Fig. 3. All the stimuli are identical except for one of the layer luminances, which increases systematically as one progresses through the series. Inspection of the stimulus set reveals that at the extreme ends of the series (top left and bottom right) the layer does not appear to form a uniform transparent surface. However, in the middle of the series one does have the percept of a milky ($F > 0.0$), semitransmissive ($0 < t < 1$), uniform transparent surface. Only one of the stimuli, however, is the correct one, at least according to the episcotister luminance model (midway in the middle row; provided that the luminances have been correctly reproduced in the figure). It is important to understand that the correct third layer luminance depends on the particular model of transparency. Although there are certain restrictions on the possible settings of the two fixed layer luminances, within those restrictions a given combination of background and layer luminances will produce a different t and F combination depending on the transparency model assumed by the visual system. In experiment 1, subjects were required to adjust this third layer luminance to provide the best perceived transparent surface for a number of stimuli with various combinations of background luminances and layer t's and F's.

The first goal of this study is to employ the six-luminance stimulus to test various models of perceived transparency and to measure both the accuracy and the precision with which observers make transparency settings with respect to their preferred model.

2. GENERAL METHODS

A. Stimulus

The stimulus consisted of two simulated concentric disks, each divided into three equal overlapping sectors, producing an illusory transparent layer on a tripartite background [see Fig. 1(c)], such that the sectors **P**, **Q**, and **S** were on top of the background portions **A, B**, and **C**. The layer luminances P and Q were calculated to produce a range of t's and F's according to the LE model. It must be emphasized, however, that this is no way precludes the subject from making settings according to another internal model. The difference in predictions between three of the models is quite small, as will be shown below; setting two of the layer luminances according to one of the models merely ensured that there was an adequate range of t and F values represented in the stimuli for any of the models. The luminance of the area surrounding the stimulus was fixed at an intensity of half the maximum luminance produced by the display (i.e., 20 cd/m^2). Stereoscopic image pairs with a maximum outer diameter of 7.5 cm $\times$ 7.5 cm were viewed through a custom-built eight-mirror stereoscope with a principal ray path length of 45 cm for a maximum visual angle of 7.1°. Even though all parts of the image were in the plane of fixation, a stereoscope was used because these experiments were part of a larger project that included some conditions where the layer was set to a different depth from that of the background. However, only the data from flat-plane stimuli are reported here.

B. Display

All experiments were performed by using an SGI (Silicon Graphics, Inc., Mountain View, Calif.) O2 workstation (150-MHz R10000 processor) with a 17-in. monitor displaying 1280 $\times$ 1024 pixels at a vertical refresh rate of 72 Hz. The luminance output of the monitor was measured by using a single-channel optometer with photometric detector (model S370, United Detector Technology) and calibrated to produce luminances between 0 and 40 cd/m^2.

C. Subjects

Three subjects, one naïve to the purpose of the experiment (KW) and the two authors (RK, FK), participated in the experiment. All three subjects were experienced psychophysical observers who had normal or corrected-to-normal vision.

3. EXPERIMENT 1: ACCURACY AND PRECISION

A. Procedure

A combination of luminances of A, B, C, and F (ranging from 0 to 30 cd/m^2), including four different values of t (0.2, 0.4, 0.6, 0.8), were preselected with the constraints that $\{A, B, C\} \neq 20\,\mathrm{cd/m^2}$ (surround color) and $\{A, B, C, P, Q, S\} < 40\,\mathrm{cd/m^2}$ (maximum luminance output). The contrasts between adjacent stimulus regions were minimized to avoid potentially extreme nonlinear behavior of the visual system. This was done by limiting the (Michelson) contrast between the background values (A, B, C) to 30%–40%. The contrast between the layer luminances is, from the model definition [see Eq. (2)], always equal to (for $F = 0$) or smaller than the contrast between background luminances. The subjects' task was to adjust the luminance S, using the computer mouse to drag a slider on the display, such that they perceived the transparent layer as a contiguous disk with uniform transmissive and reflective characteristics on the tripartite background. The luminance S was set to a random value at the beginning of each trial. An example of the range of possible settings for one stimulus is shown in Fig. 3. Each subject made three sets of adjustments on the set of 40 precomputed stimuli.

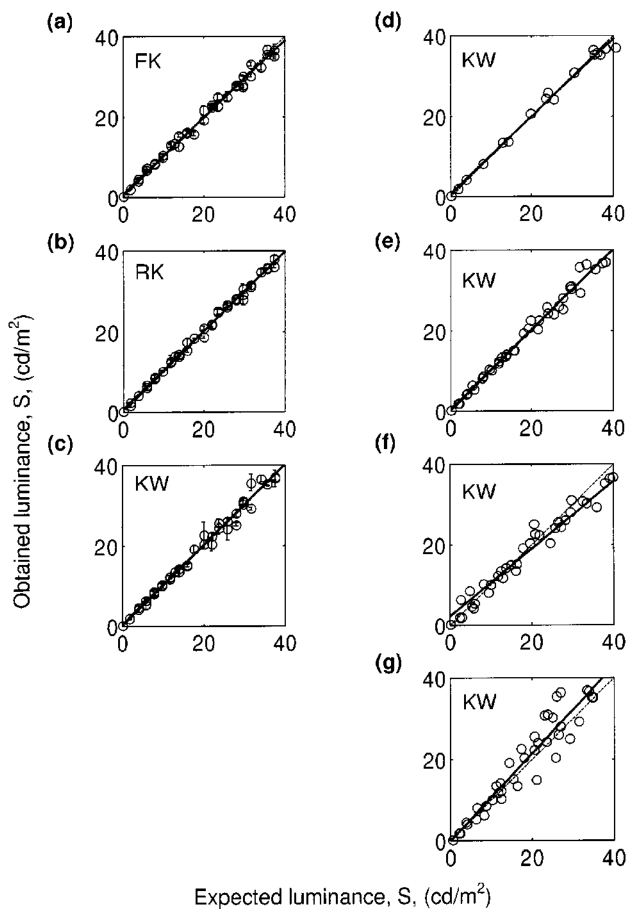

Fig. 4. The left column [(a)–(c)] shows adjustment settings for three subjects (along with solid line depicting a linear regression) compared with the expected episcotister luminance settings (dashed 45° line). The right column [(d)–(g)] shows the data from one subject (KW) compared with predictions according to alternative models: (d) filter model, (e) Michelson contrast model, (f) arithmetic mean model, (g) average brightness model. The performance of the subject (KW) matches the first three models [(c), (d), and (e)] the most closely.

B. Results and Discussion

Subjects' final settings of the luminance S were compared first with the theoretically expected luminance for each stimulus, as calculated according to the LE model: $S = tC + F$. The results for the three subjects are presented in Figs. 4(a)–4(c). Each datum point is the average of three adjustment values, with the bars indicating standard error. Points falling on or near the $y = x$ (45°) line indicate good agreement with expected values, or high accuracy. For all three subjects, the coefficient of determination (r^2) was at least 0.99. Small standard errors are the result of low variability in repeated measurements, or high precision. The results shown in Figs. 4(a)–4(c) clearly reflect very good performance to the extent that both accuracy and precision are high for all subjects.

The rest of the plots in Fig. 4 compare the results from one subject (KW) with some alternative models. The first alternative model is the filter model of Beck *et al.* Whereas the LE model does not require an illumination component (it is implicit in the additive term F), the luminance version of the filter model includes an explicit illumination variable I. For the purpose of comparison with the LE model, the highest luminance in the image was taken as the illumination component, which makes the implicit assumption that the highest luminance is white with reflectance 1.0. From the luminance formulation of the filter model,[5] we have

$$t_f = \frac{[(A - B)(P - Q)(I^2 - AQ)(I^2 - BP)]^{1/2}}{A(I^2 - BP) - B(I^2 - AQ)},$$

$$F_f = \frac{I^2(AQ - BP)}{A(I^2 - BP) - B(I^2 - AQ)},$$

$$S_f = t_f^2 \frac{I^2}{I^2 - FC} C + F_f, \tag{4}$$

where t_f and F_f are the transmittance and reflectance terms and S_f is the expected luminance setting. After stimuli that gave negative t_f and F_f values were eliminated, the expected settings for S_f given C were calculated. Even though in using Eqs. (4) one obtains different t_f and F_f values from the LE model, the resultant luminances S_f are still quite close to the LE predictions. This is not surprising, as it has been shown by Gerbino[5] that the reflectance and transmittance values of the filter model [Eqs. (4)] approach those of the LE model [Eqs. (2)] as higher and higher values of the illumination component are assumed. The fact that a number of stimuli had to be rejected because of illegal values of t_f (outside the range [0, 1]) and F_f ($F_f < 0$) suggests that the solution space of acceptable transparent percepts is smaller for the luminance filter model than for the luminance episcotister model. It should be noted that the results for the filter model apply only insofar as the assumption of the lightest patch being white is valid.

A model recently proposed by Singh and Anderson[13] employs ratios of Michelson contrasts between adjacent areas as a predictor of subjects' perceived transmittance. Instead of the luminance differences in Eqs. (3), the transmittance t_M is given by

$$t_M = \frac{P - S}{P + S} \bigg/ \frac{A - C}{A + C} = \frac{P - Q}{P + Q} \bigg/ \frac{A - B}{A + B}. \tag{5}$$

To compare this with other models, Eq. (5) was rearranged such that expected values of the luminance S could be directly plotted, giving

$$S_M = P \frac{A(AQ - BP) + C(AP - BQ)}{A(AP - BQ) + C(AQ - BP)}. \tag{6}$$

Another hypothesis is that subjects may simply have used the average contrast between the layer luminances and their background when making a setting. This would be equivalent to making the assumption that the transparent layer is a neutral-density filter, i.e., that $F \equiv 0$, even though this assumption is not supported by the physical characteristics of the stimuli. There are a number of different indices that could measure average contrast. One such index would be the arithmetic mean, i.e., $S = \frac{1}{2} C(P/A + Q/B)$. Another is the geometric mean—the equivalent of taking the average in log space—which is given by

$$\log \frac{S}{C} = \frac{1}{2}\left(\log \frac{P}{A} + \log \frac{Q}{B}\right) \text{ or } S = C\left(\frac{P}{A}\frac{Q}{B}\right)^{1/2}. \tag{7}$$

Since the geometric mean of two values is in general close to the arithmetic mean when the two samples are similar (recall that an effort was made to avoid large stimulus contrasts), in this case the two means are almost identical, and only the comparison with the arithmetic mean is shown in Fig. 4(f).

Finally, the average luminance

$$S = \frac{P + Q}{2} \tag{8}$$

was also calculated and plotted against observed data, since it is conceivable that subjects were making some sort of average brightness setting of the two fixed layer luminances, independent of the background intensities.

The results show that the LE model, the ratio of Michelson contrasts model, and, to nearly the same extent, the luminance filter model ($r^2 = 0.99$ for all three for subject KW) account very well for subjects' performance. Neither of the other plotted predictions shown in Fig. 4 matches the results as well as the LE model ($r^2 = 0.96$ and 0.90, respectively, for the arithmetic mean and average brightness hypotheses), and both the slope of the regression and the goodness of fit are inferior. Taken together, the results from experiment 1 demonstrate that with the six-luminance display, human subjects make accurate (i.e., close) and precise (i.e., reproducible) transparency judgments according to both the luminance formulation of the episcotister model and the Michelson contrast model.

The performance of the subjects in this experiment says little, however, about whether there is a range of settings around the ideal luminance S, given a set of $\{A, B, C, P, Q\}$ that nevertheless produces some degree of perceived transparency. The next experiment was designed to address this question.

4. EXPERIMENT 2: RANGE OF PERCEIVED TRANSPARENCY

To understand the purpose of this experiment, consider the following loose analogy. Imagine that one wishes to determine the wavelength perceived as unique yellow. In principle, one might adjust the wavelength of light until unique yellow was perceived. The estimate of unique yellow and the precision of the estimate would then be given by the mean and, say, the standard error of a series of such adjustments. One might then compare the measured estimate of unique yellow with a prediction based, for instance, on some physiological model of color vision. This is analogous to experiment 1, where the luminance needed to produce the best transparent percept was determined and compared with a putative physical/psychological model, and the accuracy and the precision of a range of such settings were obtained. One could, however, ask a quite different question: Over what range of wavelengths around unique yellow do subjects perceive at least some degree of yellowness? To estimate this range, we might ask subjects to indicate the wavelength boundaries on either side of unique yellow, beyond which no yellowness was perceived. This, however, would be a somewhat crude measure, since it would not take into account the *pattern* of decline of perceived yellowness as one moved away from unique yellow.

In this experiment, the range of luminances over which a degree of transparency is perceived was measured by using a novel technique that allows the determination of the pattern of decline of perceived transparency as one departs further and further from the point of ideal or best transparency.

A. Procedure

The subjects were presented with forced-choice[14] pairs of six-luminance stimuli in two temporal intervals and were asked to indicate (by a button press) which of each pair contained the more compelling transparency, insofar as it presented a layer with more uniform transmissive and reflective properties around the entire disk. Each stimulus was presented for 1250 ms, with a 250-ms interstimulus interval. When no stimulus was present, the screen was a blank mid-gray color. Whereas in the adjustment data of experiment 1 the stimuli encompassed a wide range of different combinations of parameter values, in this experiment all the trials were based on a stimulus with the fixed parameters $A = 6.0\,\mathrm{cd/m^2}$, $B = 14.7\,\mathrm{cd/m^2}$, $C = 24.0\,\mathrm{cd/m^2}$, $t = 0.75$, and $F = 9.0\,\mathrm{cd/m^2}$, resulting in expected values of P, Q, and S of 13.5, 20.0, and 27.0 $\mathrm{cd/m^2}$, respectively, according to the LE model.

In each trial two stimuli with, say, $S_1 = 24.5\ \mathrm{cd/m^2}$ and $S_2 = 34.5\ \mathrm{cd/m^2}$ were chosen from a list of nine potential luminances of S. The values of S were sampled evey $2.5\ \mathrm{cd/m^2}$, and the list was centered on the expected LE model value (i.e., $27.0\ \mathrm{cm/m^2}$), such that $S_i \in \{17.0, 19.5, 22.0, 24.5, 27.0, 29.5, 32.0, 34.5, 37.0\}$ (see Fig. 3, for example). Each stimulus pair was presented twice, in different presentation order. The angular orientation of the stimulus was randomized such that the test patch did not appear in the same location from one pair to the next, while the orientation within each pair was held constant. The trials were interleaved such that in one trial the luminances P and Q were held fixed while the luminance S was selected from an array of nine values (as above), whereas in the next trial Q and S were held fixed while the luminance P was selected from a list of nine values centered on *its* expected value, and so on. The values of A, B, C, t, and F, as well as the color of the surround, were held fixed throughout.

Three sessions of 216 trials were performed by subjects FK and KW. Each session consisted of three interleaved stimulus sets—one set for each third of the circle: **P**, **Q**, and **S**. There are $(8 \times 9)/2$ ways of selecting pairs from a list of nine. With two presentations of each pair within each of the three sessions, this yields a total of $8 \times 9 \times 3 \times 3 = 648$ trials, which were grouped together and tallied.

The subjects were, in addition, asked to make three free-adjustment settings for the three parts of the circle (as in experiment 1) to determine each subject's best-transparency setting.

B. Analysis: Modified Method of Pair Comparisons

The method of pair comparisons[15] is traditionally used to rank-order objects (e.g., carrots, zucchini, celery) in a category (e.g., vegetables) in order of preference and to subsequently find the objects' coordinate values along a unidimensional preference scale with an arbitrary origin, e.g., {−1.2, 0.2, 1.5} for {celery, carrots, zucchini}. The traditional method of pair comparisons cannot be used, however, to recover an underlying preference *distribution* where a unidimensional parameter space with equally spaced samples is already defined. In the case of this experiment, the parameter space is the luminance of the variable test patch (be it **P**, **Q**, or **S**), which is sampled strictly every $2.5\ \mathrm{cd/m^2}$. This distribution describes not only how much more compelling a stimulus is than another (by comparing the two corresponding points on the function) but also delimits the range of stimuli regarded as transparent at all. Presumably, if two stimuli are equally undesirable, each will be selected an equal number of times, and their ordinate values will be equal.

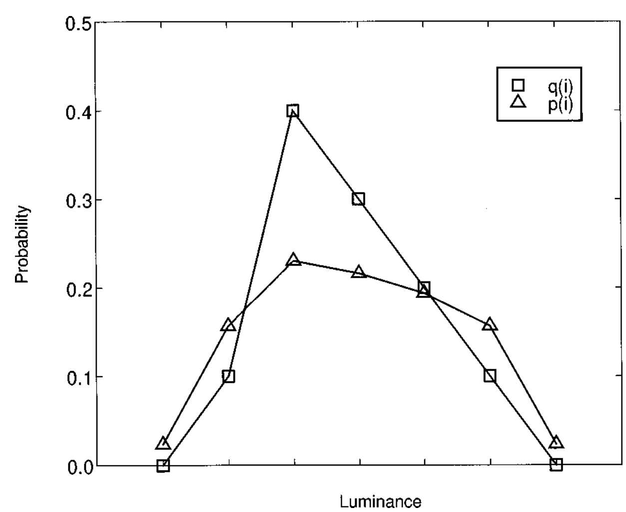

Fig. 5. Simulated example of underlying *a priori* probabilities (squares) and tally scores (triangles).

Table 1. Raw Forced-Choice Results[a]

P (cd/m^2)	FK	KW	Q (cd/m^2)	FK	KW	S (cd/m^2)	FK	KW
3.5	4	4	10.0	12	5	17.0	13	10
6.0	5	8	12.5	14	10	19.5	30	24
8.5	36	36	15.0	17	18	22.0	46	43
11.0	31	33	17.5	40	37	24.5	36	37
13.5	33	34	20.0	37	37	27.0	28	34
16.0	34	32	22.5	38	40	29.5	28	31
18.5	35	31	25.0	31	34	32.0	15	18
21.0	17	19	27.5	16	21	34.5	12	13
23.5	21	19	30.0	11	14	37.0	8	6

[a] The data for each subject are presented for each sector of the stimulus. The tallies show the number of times each stimulus was selected. Note that when the stimuli with, say, $P = 3.5$ and 21.0 cd/m^2 are compared, the luminances Q and S of the two other sectors are held fixed at the LE model values. These raw tallies were normalized with respect to the column total ($N = 216$) to give p_i for each curve in Figs. 6 and 7.

The following modification to the classical method of pair comparisons was made to recover what will hereafter be referred to as the *a priori* probability density function (PDF). It is assumed that given a range of transparency stimuli ($i \in 1, 2, \ldots, N$) that vary along a single dimension (be the relevant parameter transmittance, reflectance, or luminance of the test patch, as in this case; see Fig. 3), there exists a probability for each stimulus ($0 \leqslant q_i \leqslant 1$) to be preferentially selected over the others by an observer over a series of repeated trials. An assumption is made also that the sampled range of the parameter space is wide enough that the tails of the PDF tend to zero (that is, that the whole distribution is captured in the PDF) and that the area under the PDF is normalized, i.e., $\Sigma q_i = 1$. For example, the PDF of a series of seven stimuli might be {0.0, 0.1, 0.4, 0.3, 0.2, 0.1, 0.0}, as shown in Fig. 5. Data could also have been collected by presenting all seven stimuli arranged randomly side by side over a number of trials and asking a subject to select the best one or the one with the most consistent transparency according to some criterion, or even by doing single presentations and asking for binary (yes/no) or scalar (a score from 1 to 10) judgments, as Beck *et al.*[7,16] and others[17] have done.

To sample the parameter space by using a forced-choice paradigm with the pair comparison method, random pairs were selected from the series and the observer was asked to choose the preferred one. A tally was kept of the number of times each stimulus was selected in favor of all the others. This list of tallies (the observed data) will be called p. The *a priori* distribution sought will be labeled q, where q_i is the probability of selecting the ith stimulus over all the other stimuli in the series. Given a pair of stimuli with underlying probabilities q_j and q_k presented in a given trial, the probability of choosing stimulus j over stimulus k is simply $p_{j>k} = q_j/(q_j + q_k)$. That is, the probability of choosing q_3 over q_5 in the above example on any trial where q_3 is compared with q_5 will be 0.4/(0.4 + 0.2), or 2/3. Another interpretation would be that on average, in two out of three trials where stimuli 3 and 5 are compared, stimulus 3 would be selected. If both q_j and q_k are zero, $p_{j>k}$ is assumed to be 0.5, since the observer is forced to make a choice and will likely choose each stimulus an equal number of times. Over a series of N trials, the probability of selecting any particular stimulus becomes

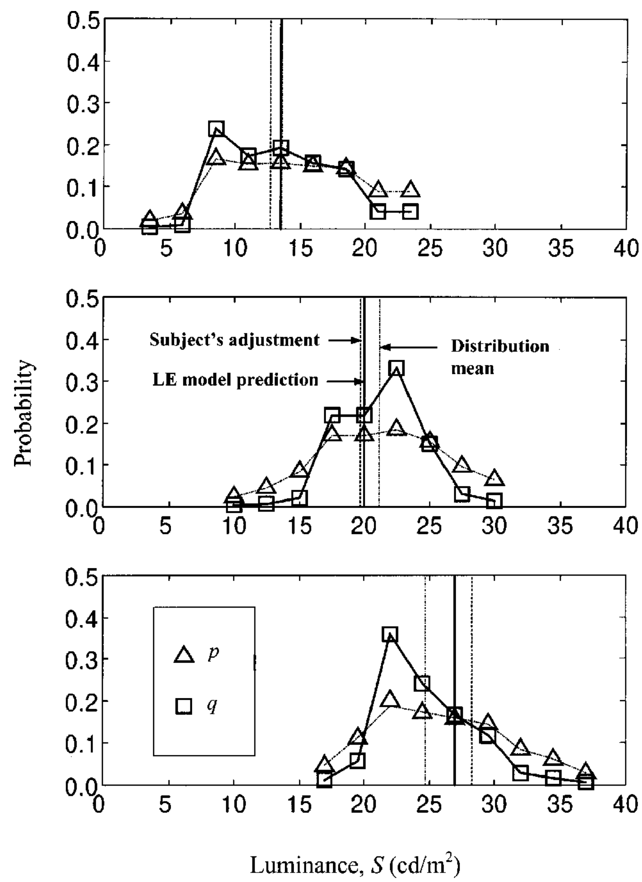

Fig. 6. Forced-choice results for subject KW. Each plot shows the results from the luminance perturbation of one part (**P**, **Q**, and **S**, respectively, from top to bottom) of the stimulus around the value according to the LE model (solid vertical line). The original tally data are shown by triangles. The squares show the recovered *a priori* probabilities q. The dotted–dashed curve joining the triangles shows a check of the solution by substituting q_i back into Eq. (9) to compare with the observed tallies p. The vertical dotted–dashed line is the first moment (average) of the *a priori* data, and the vertical dashed line is the average of three adjustment settings made by the subject.

$$p_i = \frac{1}{C_2^N} \sum_{\substack{j=1 \\ j \neq i}}^{N} \frac{q_i}{q_i + q_j}, \tag{9}$$

where

$$C_k^N = \frac{N!}{k!(N-k)!}$$

is the number of possible ways of picking k items out of N without replacement. Since one is selecting pairs, $k = 2$, and if one ensures that each pair is presented an equal number of times, $1/C_2^N$ becomes a common factor to all the terms in the sum. The terms where $j = i$ are excluded from the sum, because the same two stimuli were never compared in the same trial.

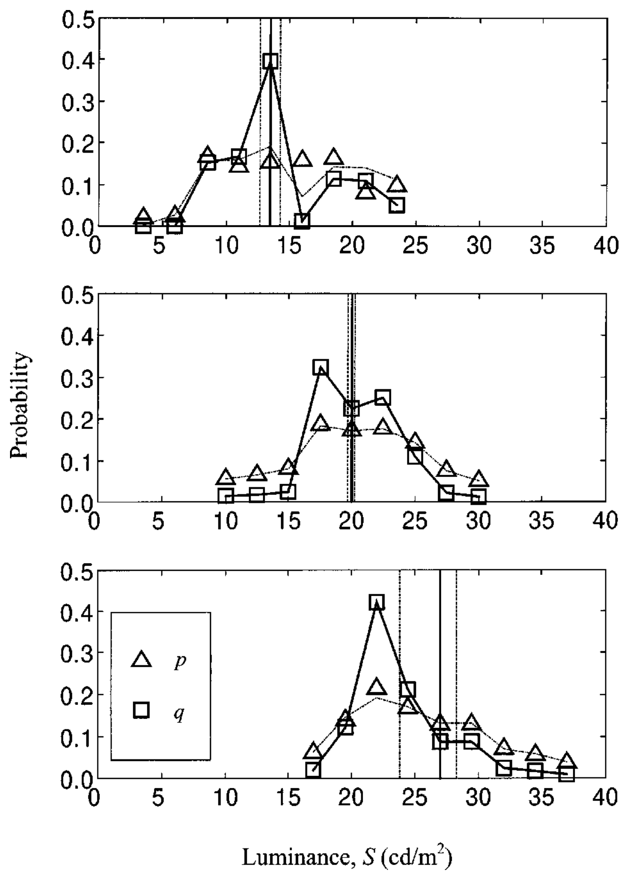

Fig. 7. Forced-choice results for subject FK.

Table 2. Summary of Forced-Choice Results[a]

	FK	KW
P, LE model value	13.5	13.5
P, free adjustment (std. err.)	16.3 (1.0)	12.7 (0.6)
P, PDF mean (std. dev.)	14.2 (6.5)	13.6 (6.5)
Q, LE model value	20.0	20.0
Q, free adjustment (std. err.)	17.9 (0.2)	19.7 (0.3)
Q, PDF mean (std. dev.)	20.3 (6.5)	21.2 (6.6)
S, LE model value	27.0	27.0
S, free adjustment (std. err.)	21.2 (0.1)	28.3 (1.5)
S, PDF mean (std. dev.)	23.8 (7.2)	24.7 (6.9)

[a] All values are luminances in cd/m^2. No distinction was made between P, Q, and S during the study, since the angular orientation of the stimulus was randomized.

The array p_i then contains the observed tallies, expressed as proportions of the number of times stimulus i was selected across all trials, and q_i are the *a priori* probabilities that one seeks to recover, as shown in Fig. 5. Since each p_i is dependent upon all the q_i (i.e., the whole PDF), inverting the transformation is not a simple task. The problem can be easily solved numerically, however, if formulated as a constrained minimization with Lagrange multipliers.[18] Specifically,

$$\sum_{i=1}^{N} \left(p_i - \frac{1}{C_2^N} \sum_{\substack{j=1 \\ j \neq i}}^{N} \frac{q_i}{q_i + q_j} \right)^2 \tag{10}$$

is minimized subject to the constraints

$$\sum_{i=1}^{N} q_i = 1, \qquad 0 \leq q_i \leq 1, \quad |q_i - q_{i+1}| < 0.3.$$

The constraints ensure that the PDF sums to unity and that each individual probability is between zero and unity. A gradient constraint was also added to avoid overshoots and undershoots in the minimization.[19] Visual inspection of the data confirms that large differences in adjacent bins are artifactual or are due to outliers.

C. Results and Discussion

Table 1 lists the raw (nonnormalized) tallies for both subjects. Figures 6 and 7 show the recovered *a priori* probabilities (q_i) as well as the raw tally data (p_i) for subjects KW and FK. The solid vertical lines show the LE model predictions. The dashed vertical lines show the subjects' average adjustment settings corresponding to each part of the stimulus, which were measured separately. Rather than imposition of a functional form (i.e., a Gaussian) on the recovered results to obtain some measure of the central tendency and spread, the average and the standard deviation of the distribution were computed. The dotted–dashed lines show the means of the three distributions. The results are summarized in Table 2.

The results show that both the subjects' free adjustments and the mean of the PDF are reasonably close to (generally within 3 cd/m^2 of) the predicted episcotister model luminances. The widths of the distributions also confirm the hypothesis that there are a series of luminances that give rise to at least some degree of phenomenal transparency. The distributions appear to be unimodal, confirming that each subject has an individual point of best transparency.

These results could also be understood in terms of balanced and unbalanced transparency. Unbalanced transparency refers to displays in which Metelli's rules regarding the limits on t and r [see Eqs. (1)] are violated. Unbalanced transparency implies the presence of two *different* transparencies with different combinations of t and r for the two different background–layer pairs, rendering Eqs. (1) indeterminate.[20] It may be that the results of this experiment reflect, in part, the preference of observers for balanced rather than unbalanced transparency.

5. GENERAL DISCUSSION

In this paper, the accuracy, the precision, and the range of perceived achromatic transparency have been investigated with the use of a new stimulus (the six-luminance display) and a new psychometric technique (the modified method of pair comparisons). In the first experiment, observers adjusted the luminance of a test patch on a putative transparency stimulus such that it created the most compelling impression of a uniform transparent disk overlying a tripartite background. The results were compared with predictions from a number of models of perceived transparency and other hypotheses concerning subject performance. It was shown that Singh and Anderson's ratio of Michelson contrast model[13] and the luminance formulation of Metelli's episcotister model[5] (in agreement with a previous study by Gerbino[10] using an eight-luminance stimulus arrangement) best predicted the subjects' adjustments. Our analysis revealed that subjects did not appear to assume that the layer was a simple neutral-density filter, reinforcing the idea that they were sensitive to the additive reflective component F. In turn, this confirms the idea that a neutral-density filter (which gives rise to multiplicative, subtractive,[21] or film[22] transparency) is perceived as a limiting case of transparency where the layer has zero reflectance.

The filter model proposed by Beck *et al.*[7] takes into account internal reflections between the filter and the background, resulting in a slightly different set of equations [Eqs. (4)]. The protocol of Beck *et al.*, however, required subjects to make a simple yes/no judgment as to whether each surface, displayed separately, was perceived as transparent. This binary response method is a cruder version of the modified pair comparison technique presented in experiment 2. Although Gerbino *et al.*[10] did not compare their results with predictions made by the filter model of Beck *et al.*, Gerbino[5] (and, presumably, Metelli[2]) has favored models compatible with the theory of color scission—where image intensities can be attributed to two superimposable sources, namely, a background and a layer. According to Metelli and Gerbino, the LE model would be preferable to the luminance filter model. Irrespective of compatibility with scission theory, the results of experiment 1 show that the LE model encompasses a wider set of perceptually valid transparency stimuli than the filter model, even though the episcotister and filter model results for the stimuli that did produce legal transmittance and reflectance values were similar. In addition, the results show that the three subjects made settings with both high accuracy and precision. This finding indicates that there is a definitive and reliable ideal-transparency point for human observers.

The second experiment explored the parameter space surrounding the predicted LE model solutions on a specific exemplar stimulus to find whether there is a range of stimuli around this predicted ideal point that also gives rise to at least some degree of perceived transparency. It was found that there is a wide range of the variable layer luminance (average standard deviation ≈ 7 cd/m^2) within which a subject would report some phenomenal transparency, even though the subject's *most preferred* point might be quite rigorously and reliably reproduced to be near the mean of this distribution.

Few studies have endeavored to determine the exact luminance or reflectance combination required to give rise to a perceptually transparent surface, whether by using a classical bipartite arrangement [such as in Fig. 1(b)] or otherwise. For example, Gerbino *et al.*[10] used two four-luminance stimuli to perform their matching experiments. The simpler and more cohesive six-luminance stimulus employed here has a number of advantages. In addition to displaying all of the relevant luminances simultaneously within central vision, it leaves the luminance of only one patch to be manipulated. This setting has only one correct value according to the subjects' internal model of transparency, since according to any algebraic transparency formulation requiring four luminances, the system's two unknown parameters can be uniquely identified, as evident from Eqs. (2). The results of experiment 2 suggest that if there is a so-called ideal point that is preferred by an observer, it is flanked by a reasonably wide range that also engenders, albeit to a lesser extent, perceptually valid impressions of transparency.

Use of the six-luminance stimulus need not be limited to the study of luminance-based cues to transparency. There are figural conditions that must be satisfied to evoke compelling transparency percepts. Variants of the six-luminance stimulus, such as those in Fig. 8, are currently being used to explore the role of X junctions. Metelli[2] provided compelling examples of stimuli where

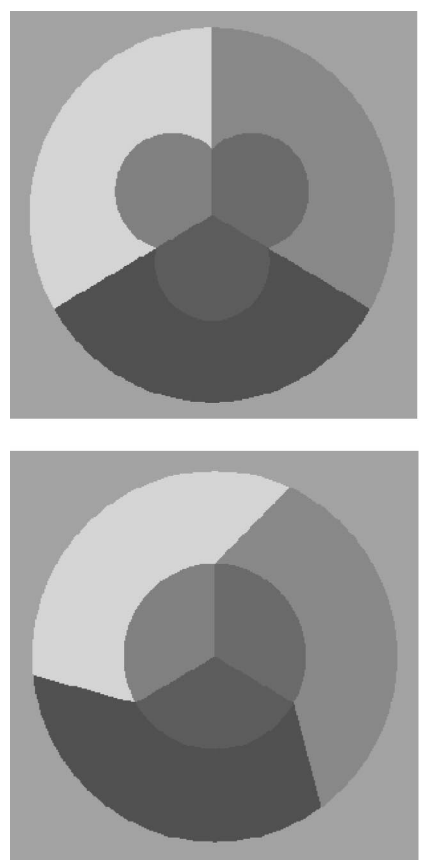

Fig. 8. Examples of modifications to the six-luminance stimulus for investigating figural conditions in transparency.

the luminances were appropriately selected but nevertheless did not produce perceptually transparent images. More recently, arguments have been made for the need for X or T junctions (the latter sometimes implying X junctions).[23] Anderson[24] and previously Beck and Ivry[16] have suggested a series of heuristics for the required contrast relationships across collinear segments of such junctions. These rules define the order of intensities (from brightest to darkest) around an X junction as a cue for transparency. These contrast polarities were satisfied as a consequence of the modified Metelli equations used in these studies; their role was not explored here. The six-luminance stimulus is ideally suited to test the importance of these various proposed figural conditions, as it is easily employed to provide a rigorous and quantitative assessment of perceived transparency rather than merely a qualitative one.

How can these findings be placed in the context of models of transparency applied to the chromatic domain? D'Zmura *et al.*[25] have proposed a chromatic transparency model that describes the characteristics of filters based on the translation and the convergence of vectors in color space. It is at present unknown how perceptual models of chromatic and achromatic transparency are related and how the conditions necessary for chromatic transparency can be collapsed from a color space into the luminance dimension. The LE model implemented here, however, corresponds to a scalar version of the vector convergence condition in the study of D'Zmura *et al.* Westland and Ripamonti[26] have also suggested a model that aims to predict cases of perceived chromatic transparency. Their studies show that ratios of retinal cone excitations (across all cone types) between a background and its corresponding transparent layer are invariant over a wide range of cases of simulated physical transparency, and they cite this invariance as a possible cue or mechanism for the detection of chromatic and achromatic transparent overlays. They are, however, careful to point out, as Metelli[2] did originally, that physical transparency does not necessarily give rise to perceptual transparency, nor do all cases of perceived transparency correspond to a physically transparent surface. In other words, the so-called proximal-to-distal mapping of perceptual versus physical transparency is not one to one but rather many to many. The real potential advantage of these newer models (such as Westland and Ripamonti's and Singh and Anderson's in the achromatic domain) over conventional models is their use of contrast measures (cone excitation ratios or Michelson contrasts) rather than differences of luminances or reflectances. Considering the wealth of literature indicating that the visual system is well adapted to detecting changes in contrast, it would be interesting to see whether the results of their simulations are supported by the type of psychophysical evidence provided by the present study.

ACKNOWLEDGMENTS

The work in this paper was supported by grant MT11554 to F. A. A. Kingdom and a doctoral research award to R. Kasrai from the Medical Research Council of Canada.

Address correspondence to Frederick A. A. Kingdom at the location on the title page or by e-mail, fkingd@po-box.mcgill.ca.

REFERENCES AND NOTES

1. H. von Helmholtz, *Treatise on Physiological Optics* (Dover, New York, 1866/1962).
2. F. Metelli, "The perception of transparency," Sci. Am. **230**(4), 90–98 (1974).
3. F. Metelli, "An algebraic development of the theory of perceptual transparency," Ergonomics **13**, 59–66 (1970).
4. The terms "layer reflectance" and "layer luminance" are used throughout this paper to denote the absolute reflectance or luminance of the stimulus corresponding to the putative transparent layer, even though the variations in luminance would, in the physical case, be due to the background.
5. W. Gerbino, "Achromatic transparency," in *Lightness, Brightness, and Transparency*, A. L. Gilchrist, ed. (Erlbaum, Hove, UK, 1994), Chap. 5, pp. 215–255.
6. A. L. Gilchrist, "Perceived lightness depends on perceived spatial arrangement," Science **195**, 185–187 (1977).
7. J. Beck, K. Prazdny, and R. Ivry, "The perception of transparency with achromatic colors," Percept. Psychophys. **35**, 407–422 (1984).
8. W. Fuchs, "Experimentelle Untersuchungen über das simultane Hintereinandersehen auf derselben Sehrichtung," Z. Psychol. **91**, 145–235 (1923).
9. S. C. Masin, "Transparent surfaces and illuminated holes," Perception **24**, 761–770 (1995).
10. W. Gerbino, C. I. Stultiens, J. M. Troost, and C. M. de Weert, "Transparent layer constancy," J. Exp. Psychol. Hum. Percept. Perform. **16**, 3–20 (1990).
11. The principle of the six-luminance method has also been described by W. Gerbino in a conference presentation, though it was not mentioned in the associated conference abstract: W. Gerbino, "Colour constraints and optimal transparency," Perception (Suppl.) **22**, 2 (1993).
12. Additionally, in this display all the relevant luminances are in close proximity in central vision, such that no scanning of the image is required.
13. M. Singh and B. L. Anderson, "Toward a perceptual theory of transparency," manuscript available from Bart Anderson, Department of Brain and Cognitive Sciences, MIT NE20-447, 77 Massachusetts Avenue, Cambridge, Massachusetts 02139; bart@psyche.mit.edu.
14. There is disagreement in the literature over the usage of the term 'forced-choice.' Some maintain it should be used only if there is a correct answer on every trial, which is not the case in this experiment. The term forced choice is meant here to imply that there are two alternatives on each trial, from which one must be chosen. See N. A. Macmillan and C. D. Creelman, *Detection Theory: A User's Guide* (Cambridge U. Press, New York, 1991) for a discussion of this issue.
15. J. P. Guilford, *Psychometric Methods* (McGraw-Hill, New York, 1954).
16. J. Beck and R. Ivry, "On the role of figural organization in perceptual transparency," Percept. Psychophys. **44**, 585–594 (1988).
17. M. Singh and D. Hoffman, "Part boundaries alter the perception of transparency," Psychol. Sci. **9**, 370–378 (1998).
18. D. G. Luenberger, *Linear and Nonlinear Programming* (Addison-Wesley, Reading, Mass., 1984).
19. The minimization was performed by using the constr() function in the MATLAB optimization toolkit (The MathWorks, Inc., Natick, Mass).
20. F. Metelli, O. Da Pos, and A. Cavedon, "Balanced and unbalanced, complete and partial transparency," Percept. Psychophys. **38**, 354–366 (1985).
21. J. Beck, "Additive and subtractive color mixture in color transparency," Percept. Psychophys. **23**, 265–267 (1978).
22. D. Kersten, "Transparency and the cooperative computa-

tion of scene attributes," in *Computational Models of Visual Processing*, M. S. Landy and J. A. Movshon, eds. (MIT Press, Cambridge, Mass., 1991), Chap. 15, pp. 209–228.
23. T. Watanabe and P. Cavanagh, "Transparent surfaces defined by implicit X junctions," Vision Res. **33**, 2339–2346 (1993).
24. B. L. Anderson, "A theory of illusory lightness and transparency in monocular and binocular images: the role of contour junctions," Perception **26**, 419–453 (1997).
25. M. D'Zmura, P. Colantoni, K. Knoblauch, and B. Laget, "Color transparency," Perception **26**, 471–492 (1997).
26. S. Westland and C. Ripamonti, "Invariant cone-excitation ratios may predict transparency," J. Opt. Soc. Am. A **17**, 255–264 (2000).