Combining achromatic and chromatic cues to transparency

Jacqueline M. Fulvio

Manish Singh

Laurence T. Maloney

We investigated how achromatic and chromatic cues interact to produce transparency. Observers were shown six-region stimulus displays similar to those used by R. Kasrai and F. A. A. Kingdom (2001) and made adjustments of the color and luminance attributes of one of the filter regions to achieve the best percept of transparency. The dependent measure of primary interest was setting reliability, the reciprocal of setting variance. We wished to determine whether the combination of chromatic and achromatic information leads to enhanced reliability of perceived transparency. In Experiment 1, we measured reliability for achromatic, \( L \), superimposed luminance with color, \( L + C \), and superimposed luminance with polarity-reversing color, \( L + iC \). We found that observers' reliability was highest for the \( L + C \) condition, consistent with effective cue combination. In a second experiment, we compared setting reliability for \( L, L + C \), and a new chromatic-only condition \( C \). In the \( L + C \) condition, observers were asked to make separate and iterative settings of luminance and color to achieve the best percept of transparency. We compared their settings in \( L \) with the luminance settings in \( L + C \) and their settings in \( C \) with their color settings in \( L + C \). Color adjustments were more reliable when accompanied by luminance information but not vice versa. In Experiment 3, we manipulated the transmittance of the achromatic and chromatic filters separately and investigated how this influences the settings made for each attribute. No systematic influence of filter transmittance on the settings made for perceived transparency was found.

Keywords: transparency, luminance, equiluminant color, cue combination

Introduction

Consider the stimulus in Figure 1A. Most readers will perceive this display as consisting of a homogeneous, partially transmissive disk overlying a three-part background, “... a transparent coloured veil... spread over the field…” (von Helmholtz, 1910/1962, pp. 285–286). A small rotation of the disk, however, disrupts the percept of transparency and reveals the three discrete colors that this region comprises (Figure 1B). The striking difference between the perceived homogenous color of the filter in the transparency display (Figure 1A) and the three sector colors comprising the central disk (Figure 1B) illustrates the phenomenon of transparency. Under appropriate image conditions, the visual system effectively assigns two sets of colors to an image region, one belonging to an underlying background surface and the other to the transparent layer through which the background is seen. This phenomenon is also referred to as “color scission” (Koffka, 1935; Metelli, 1974a).

In addition to geometric constraints (such as the continuity of contours, disrupted in Figure 1B), the percept of transparency also depends on photometric constraints. Metelli (1970, 1974a, 1974b) developed a model of achromatic transparency based upon a physical setup—an episcotister—from which he derived two qualitative constraints for the perception of transparency. This model also makes quantitative predictions for the transmittance and reflectance of a filter (such as the overlying disk in Figure 1A). The model was based on a physical setup involving a bipartite background underlying a transparent disk—thereby generating a four-region display (Figure 2A). Throughout the article, we will adopt a notation in which corresponding background and filtered regions are denoted as \( b_i \) and \( f_i \), respectively, where \( i = 1, 2, ..., n \) and \( n \) is the number of background regions. Hence, by this convention, \( n = 3 \) in Figure 1A and \( n = 2 \) for Metelli’s diagram in Figure 2A.

Given this four-region diagram, Metelli modeled the luminance \( L^* \) of the filtered regions \( f_i \) as a convex mixture of the corresponding background color \( b_i \) and a filter luminance \( t \):

\[
f_i = ab_i + (1 - a)t, \quad i = 1, ..., n,
\]

where \( a \) is the transmittance of the filter (the proportion of incident light it transmits).
When the number of background regions is \( n \geq 2 \) and perceptual decomposition of an image region yields a background surface composed of the regions \( b_i, i = 1, \ldots, n \), with an overlying transparent layer composed of the filtered regions \( f_i, i = 1, \ldots, n \), we can solve simultaneous equations based on Equation 1 to obtain the transmittance \( \alpha \) and the luminance \( t \) of the filter.

Metelli’s qualitative constraints follow immediately from the equation above, along with the assumption that the transmittance, \( \alpha \), is nonnegative with magnitude no greater than 1. The first constraint states that the region of possible transparency must preserve contrast polarity relative to the background: If a contrast border in the background is dark-light in a given direction (say, \( b_1 > b_2 \)), then its extension into the filter region must also be dark-light \( (f_1 > f_2) \). This is a powerful ordinal constraint that has been used by many researchers (Adelson & Anandan, 1990; Anderson, 1997; Beck & Ivry, 1988), sometimes with striking consequences (e.g., Anderson & Winawer, 2005). If violated, the “filter” regions are polarity reversing and transparency is not perceived. The second constraint states that the magnitude of the luminance difference within the filter region should be no greater than in the region in plain view \( |b_1 - b_2| \geq |f_1 - f_2| \). In Figures 1A and 2A, then, the conjunction of preserved contrast polarity across the boundary of the central disk and reduced contrast magnitude within it constitutes a powerful cue to transparency so that the presence of an overlying layer is perceived to be responsible for the reduced contrast.

### Transparency as convergence in color space

Metelli’s model was intended to describe transparency in achromatic displays. D’Zmura, Colantoni, Knoblauch, & Laget (1997) extended Metelli’s model to three-dimensional color space. Rather than treating \( b_i, f_i, \) and \( t \) in Equation 1 as scalars denoting luminance values, they treated them as three-dimensional vectors in color space. It follows from Equation 1 with this reinterpretation that the presence of an overlying color filter leads to a global convergence in color space. If \( b_i, i = 1, \ldots, n \), are the color vectors of the background regions and \( f_i, i = 1, \ldots, n \), are the corresponding filtered regions, the vector differences \( \Delta_i = b_i - f_i, i = 1, \ldots, n \), point to the common convergence point \( t \). The degree of convergence is the scalar \( 1 - a \), where \( a \) is the transmittance of the filter.

Metelli’s achromatic model can be expressed, then, as a special case of the convergence model where the vectors \( b_i, f_i, \) and \( t \) all lie along the luminance axis of color space and their magnitudes correspond precisely to the scalar values in his model. We will generalize Metelli’s model to the case of any three collinear vectors and their mixtures below. This will allow us to parsimoniously express conditions analogous to Metelli’s for transparency in equiluminant displays that are also consistent with the convergence model.

### Six regions for a unique solution

Although four-color displays, such as the one in Figure 2A, have been widely used in tests of the convergence model (e.g., Beck, Prazdny, & Ivry, 1984; Chen & D’Zmura, 1998; Gerbino, Stultiens, Troost, & de Weert, 1990), they have a major shortcoming. Fixing the two background luminances of the two background regions, \( b_1 \) and \( b_2 \), and the luminance of one of the filtered regions, say \( f_2 \), yields infinitely many values of \( f_1 \)
that are consistent with Metelli’s model (and the convergence model) for different values of $\alpha$ and $t$. Figure 2B depicts two such solutions for $f_1$. Because the experimental task we use involves the adjustment by observers of one of the colors in a transparency display, we naturally prefer a display configuration in which the prediction of the convergence model is uniquely defined.

As noted by Kasrai and Kingdom (2001), this uniqueness of solution is readily achieved with six-region configurations ($n = 3$ in our notation), such as the one in Figure 1A. Given fixed values for the background region luminances, $b_1$, $b_2$, and $b_3$, as well as for two of the filter region luminances, $f_2$ and $f_3$, there is a unique value $f_1$, which satisfies Metelli’s equations, consistent with the convergence model. This is demonstrated schematically in Figure 3. The four values, $b_2$ and $b_3$ and $f_2$ and $f_3$, together determine $\alpha$ and $t$. The solution for $f_1$ is now uniquely determined by these quantities.

**Combination of achromatic and chromatic cues**

In what follows, we investigate how chromatic and achromatic cues combine to determine the reliability or precision of perceived transparency. Although there has been a great deal of research on both chromatic and achromatic transparency (Beck et al., 1984; Da Pos, 1989; D’Zmura et al., 1997; Faul & Ekroll, 2002; Gerbino et al., 1990; Hagedorn & D’Zmura, 2000; Khang & Zaidi, 2002; Nakuchi, Silfsten, Parkkinen, & Ussui, 1999; Singh, 2004; Singh & Anderson, 2002), the question of how these two sources of information combine to determine percepts of transparency has not been addressed. Physical processes such as colored filters typically confound changes in luminance with changes in chroma. We will manipulate the chromatic and achromatic properties of our displays independently to avoid such confounds.

The primary measure that we will use is the reliability, $\rho$, of observers’ settings for the adjustable sector, defined as the reciprocal of setting variance (Backus & Banks, 1999). We specifically investigate whether the combination of chromatic and achromatic cues leads to a more reliable percept of transparency. For Figure 3A, we noted that in a six-region achromatic display, the convergence model predicts a unique luminance $f_1$ for the variable sector as demonstrated in the diagram in Figure 3B. In an experiment using this configuration, Kasrai and Kingdom (2001) found that although this luminance prediction for the variable sector is supported, there is in fact a relatively wide range of luminance values around $f_1$ that generate a percept of transparency. We ask whether the addition of equiluminant color to such a display will “sharpen” the percept of transparency, thereby decreasing the variability in observers’ settings.

In addition to the achromatic case—which we will refer to as the $L$ condition—we will have a purely chromatic (equiluminant color) condition, $C$. An example of such a condition is shown in Figure 4A, in which all chromatic variations are defined along the single dimension of saturation—going from neutral to highly saturated yellow. We emphasize that, in the chromatic displays, the six regions are equiluminant. It is clear that, as in the achromatic case, there is a unique solution (here, for the saturation of the adjustable sector) that satisfies the convergence model (depicted in Figure 4B). A “cue-combined” version will be created by superimposing the $L$ and $C$ displays, so that the transparent filter in the resulting $L + C$ display contains both achromatic and chromatic convergence. The details of how the superposition was achieved will be explained in the General methods section. The critical question will be whether setting reliability in the $L + C$ case is systematically higher than in the $L$ case. Moreover, to address the question of whether the introduction of any colors (rather than just polarity-preserving, transparency-consistent colors) improves the percept of transparency, we included a fourth condition, $L + iC$, in which the chromatic component of the superposition is polarity reversed—hence, inconsistent with transparency (recall Metelli’s qualitative constraints).

The first two experiments test whether adding color to luminance information makes transparency more precise relative to either cue in isolation. The third experiment...
applies further configuration manipulations to determine whether the relative strength and contrast between the two cues influence perceived transparency.

General methods

Stimuli

The stimuli were computer-generated, six-region transparency displays (after Kasrai & Kingdom, 2001) presented stereoscopically following Singh & Anderson (2002). The filter had a disparity of 19.0 arcmin. A sample stereo pair is shown in Figure 5. The three wedges (i.e., filter-background pairs) of each six-region display were randomly permuted and counterbalanced from trial to trial so that observers made an equal number of adjustments for each of the three filter regions. The adjustable filter region was always in the upper left wedge. The displays differed in their chromaticity and luminance, depending upon experimental condition.

We describe the construction of the stimuli intuitively and in detail. Imagine a globe whose north–south axis is luminance and whose equator includes the equiluminant plane of our chromatic stimuli. The North Pole, denoted $M$, is achromatic. The South Pole corresponds to the RGB value (0,0,0). We select a chromatic “yellow” point $Y$ on the equator, equiluminant to the achromatic center of the globe, denoted $B$. All achromatic regions of stimuli are weighted mixtures of $M$ and $B$. All chromatic regions are weighted mixtures of $Y$ and $B$ (and therefore equiluminant with $Y$ and with $B$). We emphasize that the convergence point is always $B$. In specifying any chromatic ($C$) or achromatic ($L$) region, we need only specify whether they are chromatic or achromatic and the weight used in Equation 1. In achromatic stimuli, for example, the weight 1 corresponds to $M$, the weight 0 to $B$.

The linearized RGB monitor values used for $M$ were (164, 164, 164), and for $B$, they were (136, 136, 136); both are neutral in appearance. We restricted ourselves to weights within the monitor gamut by choice of weights. We chose $Y$ separately for each subject to be equiluminant to $B$, based on a psychophysical procedure as discussed in the Procedure section. Note that, in terms of the convergence model, the linearized RGB vector $B$ is the convergence point $t$.

We superimposed chromatic and achromatic stimuli by vector addition of the vector differences from $B$, not vectors differences from the South Pole. In doing so, we effectively set $B$ to be the origin of our vector space and, as a result, the chromatic and achromatic components of the stimuli continued to be defined by a contraction to the point $B$. If we had instead physically superimposed the chromatic and achromatic stimuli, we would have increased the overall luminance of each region by the luminance shared by the equiluminant colors, and the convergence point would no longer be $B$. This method of superposition had an important result—superimposing chromatic stimuli on achromatic left the luminances of the achromatic regions unchanged: We could add color information without altering luminances.

To produce the six-region achromatic stimuli in Experiments 1 and 2, we first assigned each of the three background regions a weight along a 0–1 scale. The assigned weights of the three background regions $b_i, i = 1, 2, 3$, in our stimuli were always 1.0, 0.6, and 0.2, respectively. The assigned weights of the corresponding filter regions $f_i, i = 1, 2, 3$, were computed as the background weight multiplied by $\lambda = 0.4$ and were 0.4, 0.24, and 0.08, respectively. Note that we are using weights in Equation 1 to specify both the regions and the common contraction $\lambda = 0.4$ at the filter boundary. To avoid confusion in this double use of the term weights, we will

![Figure 4. Application of the convergence model to equiluminant color. The format is identical to that of Figure 3. Instead of luminance values, the model uses saturations of an equiluminant hue.](#)

![Figure 5. Example of a pair of achromatic stereoscopic stimuli used in Experiment 1. The pair on the left is for crossed fusion (\times); the pair on the right is for uncrossed (\sqcup) fusion.](#)
use a from this point onward to refer exclusively to the contraction across the filter boundary. To produce the six-region chromatic stimuli in Experiments 1 and 2, we used the same weights and the same as we did for the construction of the achromatic stimuli, but, now, we are specifying contractions from Y to B.

For Experiment 3, the chromatic component was constructed as in Experiments 1 and 2, but the achromatic component took on one of three values for : 0.55, 0.4, or 0.25. The weight computed for the adjustable filter region was randomly jittered by ±0.2 so that the initial configuration was never consistent with the convergence model.

Contrast-polarity reversal

Recall the constraints on perceived transparency articulated by Metelli: The filter must preserve contrast polarity relative to the background, and the magnitude of contrast within the filter region should be no greater than in the region in plain view. A violation of these constraints, namely, a contrast-polarity reversal caused by transposing the two filter regions, as in Figure 6A, disrupts the percept of transparency. In the six-region display of Figure 6B, the two filter regions and have been reversed, resulting in no longer preserving polarity (see Figure 6B).

In one of our conditions, the stimuli contained a polarity-reversing filter. In the superimposed luminance with polarity-reversing color condition, , the weights of the two nonadjustable filter regions of the equiluminant yellow stimulus, and , were switched after being assigned in accordance with Metelli’s equations. Then, the polarity-preserving achromatic stimulus and the polarity-reversing equiluminant stimulus were superimposed using the same method as that used for the condition.

Software and apparatus

The stimuli and experiments were programmed in MATLAB using the Psychophysical Toolbox extensions (Brainard, 1997; Pelli, 1997). The computer used in the experimental apparatus was a Sony GDM-FW 900 workstation with a 24-in. monitor displaying 1,280 × 1,024 pixels at a vertical refresh rate of 100 Hz. The color quality was 32 bit, and the graphics processor was a Quadro4 380 × 61. The system processor of the computer was an Intel Pentium 4 with SSE2. The stimuli were presented to the observers binocularly using a mirror stereoscope. Observers’ viewing position was fixed by means of a chin rest that was placed approximately 82.5 cm from the computer screen.

Task

The task of the observer was to make color and luminance adjustments of the upper left third portion of the “floating” filter until the best percept of a coherent and uniform transparent filter was evoked.

Observers

The same seven observers completed Experiments 1, 2, and 3. All were affiliated with New York University and not aware of the purpose of the experiment. All observers gave informed consent before participating in the experiment.

Procedure

Observers were first required to complete a random-dot stereogram test, in which they were to indicate whether a rectangle appeared to float in front of or behind a “background” and whether the rectangle was horizontally or vertically oriented. This test was done to ensure that all observers were able to perceive stereoscopic depth. Upon successful completion of the stereo test, observers performed a minimization-of-borders task and a minimization-of-flicker task in which they localized the yellow hue that appeared equal in intensity to the neutral gray value, . This yellow hue was then incorporated into the experimental code as the equiluminant yellow point for that particular observer.

Observers were instructed to adjust the “upper left third” portion of the filter so that it produced the best percept of a coherent and uniform transparent filter. To be certain that observers understood the notion of transparency as defined by Metelli, we provided a physical demonstration using colored gels placed over three-region paper backgrounds.
Good (polarity-preserving) cases of transparency were contrasted with bad (polarity-reversing) cases to make this demonstration. The description of how the adjustments were made and the experimental design will be described in each of the individual experiments.

Experiment 1

This experiment investigates whether the combination of achromatic and chromatic cues increases the reliability of perceived transparency relative to that of the achromatic cue in isolation. In particular, we are interested in determining whether the visual system derives a more precise percept when provided with multiple transparency-consistent signals. For this purpose, two stimulus conditions were used and the reliability of the settings for each was compared. In the achromatic L condition, the stimulus contains only gray levels (Figure 7A). In the L + C condition, the stimulus comprises a superposition of achromatic and chromatic components (Figure 7B).

As suggested by the cue combination literature, an improvement in reliability in the two-cue L + C condition indicates that perceived transparency benefits from an overall stronger sensory signal. In anticipation of finding an improvement in reliability with a combination of cues, we included a third stimulus condition, L + iC, in which an equiluminant chromatic configuration is altered to be polarity reversing prior to superposition with a polarity-preserving achromatic configuration (Figure 7C). Thus, if the reliability of perceived transparency is a function of the strength of the transparency signal elicited by the input and if a polarity-preserving chromatic configuration provides a viable signal to the transparency mechanism, then we expect the most reliable settings to be made in the L + C condition relative to the other two conditions, L and L + iC. Furthermore, the addition of a transparency-inconsistent signal as occurs in the L + iC stimulus should not provide any reliability enhancement, and therefore, we would expect that the reliability of the settings for this condition would be roughly equal to that of the L condition. If, on the other hand, the added color information improves reliability of perceived transparency but does not provide an explicit transparency signal to the visual system, then we expect that the reliability of the settings in the L + C and L + iC conditions will be equal, which in turn will be larger than that of the settings for the L condition.

Methods

Task

Observers made adjustments with a single set of keys that simultaneously controlled the weights assigned to luminance and color in the adjustable filter region. The luminance and color weights were always equal, and the settings were, thus, constrained to lie along a single line in color space.

Design

Observers were required to complete four sessions of 45 trials each, with the first session being practice. Each session contained five repetitions of 9-trial blocks, 1 trial for each of the three filter sectors per condition, yielding a total of 135 experimental trials per observer.

Results and analysis

Reliability

For each condition, the variance of the settings was determined, and the reciprocal of that value was used as a measure of reliability for that condition—greater setting reliability indicates that observers have a more stable or “sharper” percept of transparency. Having obtained these estimates, we can plot the reliabilities for each observer’s settings for one condition versus the reliabilities in another condition. If one condition consistently evokes more reliable percepts of transparency, then the data will fall on one side of the 45 deg line. Data falling on or near the
45 deg line indicate no strong difference between the reliabilities in the two conditions.

Figure 8A depicts the reliabilities of each of the seven observers’ settings, with 95% confidence intervals, for the $L + C$ condition (denoted $\rho_{L + C}$) versus the reliabilities for the $L$ condition (denoted $\rho_L$). The confidence intervals were obtained by first finding the critical values for the variance within the $\chi^2$ distribution: $\chi_{0.025, n}^2, \chi_{0.975, n}^2$, where $n = (3 \times 15) - 3$, or 42—the total number of degrees of freedom when collapsing over the three background permutations for any given condition. Using the critical values, the bounds of the confidence intervals on reliability were computed as the reciprocals of the ordinary confidence bounds on variance.

In this plot, four data points fall in the upper portion above the 45 deg line toward the $L + C$ condition, whereas three lie essentially on the line. The combination of chromatic and achromatic cues improves the reliability of perceived transparency for several observers: Performance with the combined cues is more precise than with the achromatic cue in isolation, referred to as effective cue combination (see Boyaci, Doerschner, & Maloney, 2006).

Although observers demonstrate this effect to different degrees, the results are consistent with other studies involving cue combination, which have also demonstrated individual differences in observers’ abilities to make use of multiple cues (e.g., Oruc, Maloney, & Landy, 2003). We conclude, therefore, from the general trend in the results that having both luminance and chromaticity information leads to more reliable perceived transparency.

The first plot provides evidence that perceived transparency benefits from the combination of chromatic and achromatic cues, relative to the achromatic cue in isolation. Earlier, we demonstrated the need to distinguish between true cue combination and some other beneficial use of color information. The use of the $L + iC$ stimulus configuration provides us with insight regarding this distinction. Figures 8B and 8C contain the comparisons for this condition.

In Figure 8B, the reliability $\rho_{L + C}$ of the settings for the $L + C$ condition is plotted versus the reliability $\rho_{L + iC}$ of the settings for the $L + iC$ condition. The data again fall largely in the upper portion of the plot, indicating that reliability is greater when added color is consistent with transparency (i.e., polarity preserving).

Figure 8C provides additional insights. Here, the reliability $\rho_{L + iC}$ of the settings for the $L + iC$ condition is plotted versus the reliability $\rho_L$ of the settings for the $L$ condition. Most observers actually experience a decline in the reliability of perceived transparency when added color is inconsistent with transparency, relative to the achromatic condition. This demonstrates the sensitive nature of the transparency mechanism to the strength of transparency-consistent sensory signals.

This is also an indication of nonindependent processing of luminance and chromaticity information in the derivation of perceived transparency. If the mechanisms responsible for deriving transparency maintained independent processing of the two sources of information and only incorporated the outcomes of these processes at the final stage, then we would expect that the reliability of the settings for $L + iC$ condition would be as reliable as the settings for the $L$ condition because the transparency-inconsistent color information would not provide anything useful for the resulting transparency percept and the luminance processing would become the source of information upon which the resulting percept is based.
However, for most observers, there is a decline in the overall reliability of the settings for the $L + iC$ condition relative to that of the $L$ condition, indicating that both sources of information interact to produce the resulting percept.

In sum, we conclude from the results of the reliability analyses of Experiment 1 that the precision of perceived transparency benefits from a combination of transparency-consistent chromatic and achromatic cues relative to the achromatic cue in isolation.

**Mean settings**

Although setting reliability is our primary measure of performance, we may also investigate whether the benefit of the combined-cue condition, $L + C$, is also reflected in the mean settings. Recall that our stimuli were constructed using a convergence model yielding a unique solution for each configuration that is polarity preserving. Kasrai & Kingdom (2001) found in their study that observers’ mean settings are highly accurate with respect to the predicted ones of this model under the conditions of the achromatic six-region configuration. Analyzing the mean settings obtained from the seven naive observers in our study for both the $L$ and $L + C$ configurations, we find that all but one observer also conform well to the predictions of the convergence model under both stimulus conditions—there is no effect of combined cues on the mean settings. The two plots in Figure 9 depict the mean settings for each of the seven observers in the $L$ (denoted $W_L$ in Figure 9A) and $L + C$ conditions (denoted $W_{L+C}$ in Figure 9B), respectively. The anomalous seventh observer’s data (represented by black triangles surrounded by dashed circles in both plots) reflect a different strategy in making the settings. Whereas the other six are clearly deriving accurate estimates of $a$ and $t$ and making their adjustments in accordance with those estimates (as evidenced by an increase in the adjusted weight with increasing weight of the background region, in agreement with the predictions), the anomalous observer’s strategy appears to be one of uniformity within the entire three-sector filter region (indicated by that observer’s settings being nearly equal for each of the three filter sectors).

In Figure 10A, we have isolated the anomalous observer’s data set from the other six “converging” observers and collapsed across them for the $L$ and $L + C$ conditions, which have the same predicted settings. Results for the anomalous observer are reported separately in an online supplement (Figure S1). The mean settings are highly accurate with respect to the predicted ones ($\text{RMS}_{L} = 0.0241$ [$L$]; $\text{RMS}_{L+C} = 0.04$ [$L + C$]), thereby replicating Kasrai & Kingdom’s (2001) accuracy results for achromatic (black markers) configurations and extending them to chromatic (green markers) configurations. In Figure 10B, we have replotted the collapsed mean adjustments in green for the six converging observers in terms of the convergence diagram for the six-region display introduced in Figure 3. As expected, they match well with the diagram, with only a slightly worse performance in the $L + C$ condition.

For the anomalous observer’s mean settings, see Figure S1 of the online supplement. It is interesting to note that the reliability of the settings is also greater in the combined-cue condition than in the isolated-cue condition for this observer’s strategy.

**Discussion**

While most observers adhere to the convergence model for both the $L$ and $L + C$ conditions, they do differ in terms of the precision with which they do so. In particular, the reliability of perceived transparency increases with the combination of multiple transparency-consistent (i.e., polarity-preserving) cues—in other words, we have demonstrated effective cue combination for transparency. Whereas Kasrai & Kingdom (2001) found that observers made a fairly wide range of settings around the optimal one, our findings imply that the number of “acceptable” settings decreases when a greater amount of transparency-consistent information is available.

We have not, however, tested for optimal cue combination. Because we do not have estimates of observers’

![Figure 9. Experiment 1: Mean settings (W). Mean settings for each of the three filter sectors for each of the seven observers plotted among the settings predicted by the convergence model (blue line). (A) Settings made in the L condition. (B) Settings made in the L + C condition. The settings for one anomalous observer are marked with red dashed circles (see text).](image)
reliability for an equiluminant color-only configuration, we are unable to determine whether the reliability of the $L + C$ condition is an optimal combination of the two sources of information. We will address this in Experiment 2.

**Experiment 2**

The results of Experiment 1 demonstrated that the combination of two sensory signals consistent with transparency, chromatic and achromatic, increases the reliability of perceived transparency. When observers made their settings in Experiment 1, their adjustments were constrained to fall along a single line in color space, such that the luminance and chromatic weights of the adjustable filter portion were constrained to be equal. Although the stimuli were rendered such that the convergence model predicted the same adjusted weights for both cues, imposing this constraint does not allow one to test whether observers would indeed set the two to be equal if permitted to set them separately or to make more precise statements regarding how the two cues actually combine to yield perceived transparency.

In Experiment 2, we no longer imposed the single-axis constraint on the settings and allowed observers to freely make separate and iterative adjustments of luminance and color. The stimuli were again rendered to have the same predicted solution for both cues, but by allowing the observers to adjust the chromatic and achromatic attributes of the filter portion separately, we can determine whether they are accurate along both dimensions or perhaps tend to favor one cue over the other. Additionally, we can investigate the nature of the increase in reliability of perceived transparency by comparing the reliabilities among the separate settings. We may look for trends such as the reliability for both cues increasing equally, both increasing but to different extents, one showing a substantial increase whereas the other remains constant, and so forth. Lastly, we incorporated an equiluminant yellow stimulus, $C$ stimulus (Figure 7D), thereby yielding three stimulus types for this experiment: $L$, $C$, and $L + C$. We now have the opportunity of testing for optimal cue combination of the $L$ and $C$ information within the $L + C$ context, which we could not do previously in Experiment 1.

**Methods**

**Task**

The observers were provided with two sets of keys—one controlling the luminance of the filter and one controlling the saturation of the filter. In the superimposed condition, observers used both sets in any order or combination they preferred until the filter again evoked the best percept of transparency. In the single-condition cases, observers only made adjustments with the appropriate set of keys, as the second set did not alter the display in any way. In the $L + C$ conditions, we will have separate estimates of reliability for the two settings, denoted $p_{L+C}$ (luminance) and $p_{C}$ (chromatic), and two separate mean settings, $W_{L+C}$ and $W_{C}$.

**Design**

Observers were required to complete four sessions of 45 trials each, with the first session being practice. Each
session contained five repetitions of 9-trial blocks, 1 trial for each of the three filter sectors for each of the three conditions, yielding a total of 135 experimental trials per observer.

Results and analysis

Reliability

As in the previous experiment, the reliability of the settings in each condition was computed as the reciprocal of the variance, with 95% confidence intervals. The first comparison is depicted in Figure 11A. Here, the reliability \( \rho_{C}^{L+C} \) of each observer’s color adjustments in the \( L + C \) condition is plotted against the reliability \( \rho_{C} \) of the color adjustments in the \( C \) condition. As before, the data fall toward the upper portion of the plot, indicating that the reliability of the color adjustments increases in the presence of luminance information.

Figure 11B shows the reliability \( \rho_{L}^{L+C} \) of the luminance adjustments in the \( L + C \) condition versus the reliability \( \rho_{L} \) of the luminance adjustments in the \( L \) condition. Here, the results show that there is no systematic increase in reliability for the luminance adjustments with the addition of color. Instead, for most observers, the luminance adjustments are more reliable when there is no chromatic information available. This finding appears to contradict the conclusion that the reliability of perceived transparency benefits from the combination of achromatic and chromatic cues because only the perceived transparency in the color component of the stimulus improves in the combined-cue condition. Moreover, the asymmetry in the results precludes any test for optimal cue combination because this requires that both settings are more reliable in the presence of the other cue and that the reliability in the combined conditions reflects a weighted combination of the reliabilities of the two cues in isolation.

Next, we analyze the mean settings in each of these conditions for any trends not apparent from the reliability analysis to gain insight into the cause of the asymmetry we find.

Mean settings

The mean settings obtained in Experiment 1 indicated that observers who adopt the strategy of the convergence model are highly accurate for both the \( L \) and the \( L + C \)
conditions with respect to the model’s predictions—the benefit of the combined-cue condition was seen in the reliability of the settings but not in their accuracy. This implies that the mean settings for the separate color and luminance adjustments of Experiment 2 should also be accurate both in isolation and in the presence of the other cue. However, recall that, in the first experiment, observers’ adjustments were constrained so that both the color and luminance attributes within the adjustable sector were always equal. Hence, given that there is an asymmetry in the effects of the combined-cue condition on the reliability of the independent settings, we will investigate how the mean settings compare among the isolated and combined-cue conditions, as well as how they compare among luminance and color settings.

The two plots in Figure 12 contain the collapsed mean luminance (Figure 12A) and color (Figure 12B) settings for the six converging observers. Within each graph, the settings made in the combined condition ($W_{L+}C$, $W_{C+}L$) are plotted in green, while the settings made in the isolated conditions ($W_L$, $W_C$) are plotted in black. The luminance settings $W_L$ (in isolation) and $W_{L+}C$ (combined) appear much like the mean settings from Experiment 1: highly accurate and consistent among the isolated and combined-cue conditions (RMS = 0.0195 [L]; RMS = 0.0297 [L of $L+C$]). The color settings, however, are far less accurate (RMS = 0.0639 [C]; RMS = 0.0740 [C of $L+C$]) and do not conform well to the convergence model. In particular, the mean settings $W_C$ and $W_{C+}L$ made over the more saturated backgrounds (0.6 and 1.0 in weight) are nearly equal.

Let us now consider the collapsed mean adjustments replotted in green in the convergence model diagrams in Figure 13. In both conditions, the luminance adjustments maintain the high degree of accuracy noted in the first experiment. The color adjustments, however, exhibited a systematic pattern of deviation: They converged too much (i.e., toward uniformity) and did so toward a filter color more saturated than the one predicted by the convergence model. Thus, observers’ settings of color are far from being veridical with respect to the convergence model as compared with their settings of luminance.

The mean settings for the anomalous observer are plotted separately in Figure S2 of the online supplement. Again, there is a clear strategy toward uniformity in all four conditions. Additionally, this observer showed an increase

Figure 13. Experiment 2: Convergence model diagrams of the collapsed mean settings. (A) Mean luminance settings of the $L$ condition. (B) Mean color settings of the $C$ condition. (C) Mean luminance settings of the $L+C$ condition. (D) Mean color settings of the $L+C$ condition.

Figure 14. Experiment 3: Reliability ($\rho$) data comparisons with 95% confidence intervals. (A) Reliability of the color settings made in each of the $La+C$ conditions versus the reliability of the color settings made in the $C$ condition. (B) Reliability of the luminance settings made in each of the $La+C$ conditions versus the reliability of the luminance settings made in each of the $La$ conditions. Red symbols: $\alpha_L = 0.25$; blue symbols: $\alpha_L = 0.4$; black symbols: $\alpha_L = 0.55$. 

Downloaded from jov.arvojournals.org on 05/12/2019
Discussion

In apparent contradiction to the results of Experiment 1, the reliability analysis of Experiment 2 found that although color settings are more reliable in the combined-cue condition, luminance settings are more reliable in the achromatic display than in the superimposed display. This asymmetry can be understood in terms of the mean settings, however, which remained accurate for the luminance settings in both conditions but deviated in a patterned way for the color settings. We suggest that because the color settings deviate from the model predictions, they make the luminance adjustment more difficult and, therefore, less reliable in the combined-cue case than in the luminance-only condition. By contrast, in Experiment 1, because both attributes were constrained to vary together, the accuracy of both cues could be achieved, and hence, a greater reliability was obtained in the combined-cue condition than in the isolated-cue condition.

At this stage, we may also address the anomalous observer’s unique performance in Experiment 2 in which the reliability of both cues in the presence of the other one (this observer’s data are marked by the presence of a red asterisk in Figure 11).

Experiment 3

An asymmetry in the reliability results of Experiment 2 prompted us to scrutinize the mean settings for an explanation. This investigation led us to identify a patterned deviation from the convergence model in the color settings. Specifically, these settings tended toward being too saturated and too similar among the three filter sectors than would be predicted by the convergence model. This raises the question of whether contrast within the equiluminant yellow configuration is playing an unforeseen role. Although the filter transmittance, $\alpha$, was equated for the achromatic and chromatic configurations, the apparent contrast may not have been large enough in the equiluminant yellow configuration, relative to that of the achromatic configuration, and therefore, the overly saturated settings could have resulted from attempts to “boost” the boundary contrast of the color component of the stimulus. Then, as described above, this overcompensation...
in the color settings creates a patterned deviation from the convergence model that in turn disrupts the reliability of the luminance adjustments. Experiment 3 tests this hypothesis by asking whether contrast is in fact responsible for the results obtained in Experiment 2, as well as whether those results can be altered through a manipulation of the relative filter-background contrast among the color and luminance attributes of the superimposed stimuli.

To do so, we replicated Experiment 2, but this time, we generated the stimuli so that the chromatic and achromatic configurations had different degrees of transmittance within the same display. Specifically, the filter transmittance $\alpha$ was held constant in the chromatic component but varied in the achromatic component—with one value greater than, one equal to, and one less than the chromatic transmittance in the equivalent condition in Experiment 2.

**Methods**

**Stimuli**

The stimuli were the same as those used in Experiment 2; however, the filter transmittance, $\alpha$, was manipulated. The $\alpha$ used in the chromatic configurations was held constant at a value of 0.4, whereas the achromatic configurations could take on one of three $\alpha$ values: 0.55, 0.4, or 0.25.

**Task**

The task of the observers was the same as that of Experiment 2.

**Design**

Observers were required to complete six sessions of 42 trials, again with the first being practice. Each session contained two repetitions of 21-trial blocks, 1 trial for each of the three filter sectors for each of the seven conditions, for a total of 210 trials per observer.

**Results and analysis**

**Reliability**

In Figure 14, the reliability comparisons are once again plotted for the color adjustments of the $L\alpha + C$ condition $\rho_{L\alpha + C}$ versus $\rho_{C}$ of the $C$ condition (Figure 14A) and for the luminance adjustments $\rho_{L\alpha + C}$ of the $L\alpha + C$ condition versus $\rho_{L\alpha}$ of the $L\alpha$ condition (Figure 14B). Within each plot, the reliabilities for the three luminance $\alpha$ values are plotted such that data obtained for the $\alpha_L = 0.25$, 0.4, and 0.55 conditions are in red, blue, and black, respectively.

We have clearly replicated the asymmetry found in Experiment 2, whereby the reliability of the color settings is greater when luminance information is available, whereas the reliability of the luminance settings is greater when luminance information is presented in isolation than in combination with color. In this experiment, the conditions where $\alpha = 0.4$ are identical to corresponding conditions in Experiment 2. However, the asymmetries in reliability are much more pronounced. We have no explanation for this discrepancy. Nevertheless, we now add to our findings the result that the transmittance of the filter in each configuration, $\alpha$, does not influence the surprising pattern of results. This is indicated by the fact that the graphs do not contain three distinct clusters, with one corresponding to each $\alpha$ condition. This finding makes it unlikely that a difference in apparent contrast is responsible for the asymmetry found in the results of Experiments 2 and 3.

**Mean settings**

In Experiment 2, we found an analysis of the mean settings to be instructive in reconciling the asymmetry in the reliability results. Although there is no evidence that $\alpha$ systematically influences the reliability of the settings, we would like to determine whether $\alpha$ governs the choice of setting. Moreover, we would like to further investigate the suggestion that the asymmetry in reliability is a consequence of relative contrast within the stimulus configurations.

Figure 15 contains plots of the mean luminance settings relative to the predicted settings (solid line) for the six converging observers. Each column corresponds to one of the three $\alpha$ conditions. Note that a larger value of $\alpha$ corresponds to a more transmissive filter and, therefore, one with less contrast relative to the background regions. Within each graph, the black symbols correspond to the
data for the settings made in the $L$ condition, and the green symbols correspond to the data for the luminance settings made in the $La + C$ conditions, where the filter transmittance for the chromatic configuration is held constant. While performance remains highly accurate for the six observers for all three $a$ conditions (see Table 1), a trend is evident. The mean settings tend to become more accurate with decreasing $a$ or increasing contrast between the filter and the background. This is particularly the case for the luminance adjustments made in the $L$ conditions (e.g., note the accuracy difference between the $a = 0.55$ and $a = 0.25$ conditions).

Figure S3 of the online supplement contains the mean settings for the anomalous observer. They demonstrate the same pattern of settings for each of the three $a$ conditions (essentially identical settings that are independent of background color).

Figure 16 contains plots of the mean color settings relative to the predicted settings. The left-hand graph shows the color settings for the $C$ condition. The collapsed settings for the six observers are shown in black symbols, while the settings for the anomalous observer are shown in green symbols. The six observers are less accurate (RMS = 0.0457) in this condition than they are when making the luminance settings; however, they still show a systematic increase in mean setting as predicted by the convergence model. The right-hand graph contains the color settings made for each of the three $La + C$ conditions, where $a_C$ is always equal to 0.4, but $a_L$ varies. The data for all seven observers have been collapsed for this graph, given that, here, even the converging observers no longer make their settings in accordance with the convergence model but, instead, in accordance with a strategy of uniformity among the filter regions. We saw evidence of a compromise between the convergence model and the uniformity strategies in the performance of mean color settings in Experiment 2, but in Experiment 3, uniformity has become the prevailing strategy. There is no effect of $a$, however, as all color settings are set to be highly saturated and similar.

Discussion

We motivated Experiment 3 by posing the question of whether relative contrast of the achromatic and chromatic filter regions with respect to their backgrounds contributed to the asymmetry in the reliability results of Experiment 2. We considered in particular whether the overly saturated settings of color were due to an attempt to overcompensate for lesser contrast between the filter and the background, relative to the contrast between the achromatic filter and its background in the $La + C$ stimulus. This was investigated in Experiment 3 through the use of superimposed stimuli containing varying degrees of filter transmittance, $a$, among the chromatic and achromatic configurations.

The asymmetry in the reliability of the settings was replicated in all conditions including when the transmittance of the luminance filter was greater than that of the color filter, thereby causing it to have less contrast relative to its background. A consideration of the mean settings again revealed accurate performance for the luminance settings both in isolation and in the presence of color, with a slight decline in performance with increasing filter transmittance (i.e., decreasing contrast between the filter and background sectors).

The color settings, on the other hand, were less accurate; nevertheless, they were still in accordance with the convergence model for the $C$ stimulus, but for all seven observers, they were set to be nearly equal for each of the three $La + C$ conditions—specifically, they were all highly saturated and uniform. This makes it unlikely that setting reliability and accuracy and the asymmetry in the reliability results are a function of relative stimulus contrast, as determined by the filter transmittances of the two configurations.

General discussion

Perhaps the most relevant previous study on the combination of chromatic and achromatic information is that of Kingdom, Beauce, & Hunter (2004). In a forced choice between two displays depicting possible shadows, Kingdom et al. asked observers to identify the display with the “correct shadow.” They found superior performance in displays with chromatically variegated backgrounds as compared with achromatic backgrounds. They proposed that color can help to disambiguate shadows from reflectance changes via a simple rule: A luminance change that is accompanied by a color change is a reflectance boundary, whereas a luminance change unaccompanied by a color change is a shadow boundary.

Kingdom et al.’s rule relies on the assumption that shadows generally produce only an achromatic shift in color space—in particular, a uniform darkening or convergence toward black. In this respect, the context of transparency is substantially more general than shadows because the photometric transformation introduced by the presence of a transparent filter generally has both a chromatic and an achromatic component. As we have seen in the convergence model above, the photometric convergence introduced by transparency can be directed toward any point in color space, depending on the color $t$ of the filter.

Given that reflectance boundaries and transparency boundaries can both generate chromatic and achromatic shifts in color space, it is evident that the rule for shadows articulated above cannot distinguish between reflectance changes and photometric shifts due to transparency. Thus, the presence of color, which allows for the identification of
shadows using a simple rule, does not similarly allow for the identification of transparency.

Accordingly, we did not constrain observers’ settings or provide feedback to observers based on any model of transparency. We simply asked them to make the display as consistent with homogeneous transparency as possible and assessed their reliability in doing so. The results of the first experiment indicate that the reliability of transparency percepts is greater when two transparency-consistent cues combine in a single display, although combining cues does not influence the mean settings. It is interesting to note that recent psychophysical work has tested Metelli’s quantitative predictions and found failures. In particular, the solution for \( a \) fails in general to predict perceived transmittance (Robilotto, Zhang, & Zaidi, 2002; Singh & Anderson, 2002), despite the fact that Equation 1 does in fact capture the lightness of surfaces seen through transparency (Singh, 2004). In contrast, we found that the quantitative predictions of the convergence model (and thus Metelli’s model) proved accurate for the six-region display, as all but one observer made adjustments in accordance with the model. It is normal to expect that a strategy based on equating the ratio of contrasts and one based on equating Metelli’s \( a \) (i.e., ratio of luminance differences) should result in different predictions. It turns out, however, that in the context of six-region displays, the two strategies lead mathematically to the same predicted setting for the third inner sector (see Singh, 2004). There is thus a principled reason behind why Kasrai & Kingdom (2001) obtained equally good fits to their data with the two models.

In Experiments 2 and 3, six of seven naive observers adopted the strategy of the convergence model. The initial cue-combination-like findings of Experiment 1 were challenged, however, by an asymmetry in the reliability of the settings: The color settings remained more reliable in the combined-cue conditions, whereas the luminance settings were more reliable in the achromatic single-cue conditions. The mean settings for these observers revealed highly accurate performance in the luminance settings but poor performance (with respect to the convergence model) in the color settings, particularly in the combined-cue conditions. One anomalous observer, however, adopted a strategy of uniformity within the three sectors of the filter region. This observer showed an increase in setting reliability for both cues in the combined-cue display. The mean settings for this observer indicated a single adjustment strategy for both luminance and color settings.

A single explanation can account for these two sets of results: The reliability of perceived transparency is greater when the settings for each of the cues are made in accordance with an observer’s own internal model of transparency. This follows suit with cue combination, as multiple redundant sensory signals yield more reliable percepts than singular ones.

It is puzzling, though, that the six observers who adopted the strategy of the convergence model for all luminance settings and for the color settings of the \( C \) displays were unable to produce accurate color settings for the \( L + C \) and \( L a + C \) conditions of Experiments 2 and 3. Manipulating the relative filter-background contrast for the chromatic and achromatic configurations did not alter performance, indicating that the highly saturated settings are not the result of an attempt to boost the filter-background contrast in the color portion of the stimulus. Instead, this may simply reflect an inherent weakness in the ability of the mechanisms responsible for deriving transparency percepts to use color information when luminance information is present.

**Acknowledgments**

We would like to thank Hadley Tassinari for help in constructing the apparatus, Todd E. Hudson for useful suggestions, and Michael S. Landy for comments on a previous draft. This research was funded in part by Grant EY08266 from the National Institutes of Health (L.T.M.) and Grant BCS-0216944 from the National Science Foundation (M.S.).

Commercial relationships: none.

Corresponding author: Jacqueline M. Fulvio.
Email: jmf384@nyu.edu.
Address: Psychology Department, New York University, 6 Washington Place, Room 877C, New York, NY 10003, USA.

**Footnotes**

1In Metelli’s original formulation, the model was expressed in terms of reflectance values, that is, with the filter properties being its reflectance and transmittance. However, Gerbino et al. (1990) have shown that the same equations are obtained in the luminance domain, under the assumption that all surfaces are uniformly illuminated.

2We noted in pilot studies that subjects did not always group the stimuli as a central disk with surrounding background when the stimulus was confined to a single depth plane. By presenting the stimuli in stereo, we force the visual system to segment the scene in accordance with the correct grouping we intended—the central region is perceived as separated from the background. Subjects are then left with the task of making the central region as uniformly transparent as possible through adjustments of one of the filter regions.

3Because a gel itself cannot be polarity reversing, the “bad” cases of transparency were made by altering the color of the paper background that rested under the gel, which produced the effect of a polarity reversal and, therefore, a loss of perceived transparency.
Here, RMS is the square root of the average of the squared deviations from those predicted by the convergence model: \(\sqrt{\frac{1}{n} \sum_{i=1}^{n} (X_i - X_{predicted})^2 / n}\).

Either of the following two explanations could account for this: (1) With increasing filter transmittance, the “true” settings for the filter regions are more displaced from each other (i.e., less uniform), which requires that observers make a wider range of settings, so that by virtue of this, there is more room for error. (2) Given that the color settings are always improperly set (see next paragraph in the text), having a smaller degree of filter-background contrast in the luminance configuration will only compound the difficulty in making the adjustment and contribute to a greater magnitude of error.

References


