Perceptual preferences in depth stratification of transparent layers: Photometric and non-photometric factors

Franco Delogu
Laboratory for Perceptual Dynamics, RIKEN Brain Science Institute, Wako-Shi, Saitama, Japan

George Fedorov
Laboratory for Perceptual Dynamics, RIKEN Brain Science Institute, Wako-Shi, Saitama, Japan

Marta Olivetti Belardinelli
Department of Psychology, Sapienza University of Rome, Rome, Italy

Cees van Leeuwen
Laboratory for Perceptual Dynamics, RIKEN Brain Science Institute, Wako-Shi, Saitama, Japan

In three experiments, using a two-alternative forced-choice task, we obtained depth judgments of displays containing transparent regions. The regions varied in lightness, size, and animation. Observers nearly always strongly preferred one certain depth ordering among the regions, even though their lightness conditions were expected to give rise to ambiguity among possible orderings. This expectation was based on the contrast polarity model, which expects ambiguity in the absence of contrast polarity reversal. The expectation was founded also on a stronger condition based on the transmittance anchoring principle, which gives preference to the largest lightness contrast between regions. In the absence of contrast polarity reversal and in conditions of balanced regional contrast, preferences were shown to depend on additional conditions of contrast between two respective regions and their overlap. Depth ordering judgment seems to be based on a critical decision threshold, independently of the coordinate system used to specify lightness. We also investigated the role of non-photometric factors such as motion and relative size, and concluded that these variables can modulate depth ordering judgments in transparency.

Keywords: achromatic contrast, lightness, motion, depth, shape


Introduction

The phenomenon of transparency reminds us that perception is more than what meets the eye. The experience that one object is visible through another implies that the observer is aware of two surfaces at different depths in the same location of the visual field. From a single luminance value, the visual system has the task to disentangle the contributions of different layers located at different depths. This process was known as scission to the Gestaltists (Koffka, 1935), who studied it extensively.

As an underlying object remains visible through a transparent occluder, transparency can be used by the visual system as a cue in depth perception. Moreover, scission could, in principle, be performed in many different ways. Nevertheless, perceivers are often in one mind about the preferred depth stratification of a transparent pattern. Several authors have described this remarkable consistency in terms of photometric factors (Adelson & Anandan, 1990; Anderson, 1997, 2003; Beck & Ivry, 1988; Metelli, 1974, 1985). This study investigates the question whether photometric laws are sufficient to make a given depth order of transparent layers more likely to be perceived. A particular consequence of the sufficiency assumption is that for those patterns where the photometric laws fail to specify a preference, ambiguity should arise. Our present aim is to test this prediction. Rather than to ask our subjects for transparency judgments, which may evoke implicit or explicit knowledge about the psychophysics of transparency, we simply present transparent displays and ask observers which of two possible surfaces is in front or behind.

Historically, the most influential model of transparency was based upon a specific physical configuration called the episcotister (Metelli, 1974, 1985). This is a disk with an open sector that, when rotated at high speed on a
background divided into two halves of different luminance (or color), elicits the illusory impression of transparency. Through the episcotister it is possible to derive equations for transmittance and reflectance of the transparent layer and to formulate the necessary conditions for the emergence of perceptual transparency. These conditions state that a transparent layer cannot reverse the polarity of an underlying contour (polarity constraint), and that the transparent layer must reduce the luminance difference of the underlying contour (magnitude constraint). As Metelli was well aware of, however, the model is restricted to the specific episcotister constellation; this excludes from analysis, for instance, the perception of scenes composed of multiple transparent regions and the representation of multiple depth planes.

Later studies have addressed the issue of how observers perceptually arrange the depth order of multiple, possibly transparent layers (Adelson & Anandan, 1990; Anderson, 1997, 2003; Beck & Ivry, 1988). In these studies, the notion of contrast polarity gained central ground (Adelson & Anandan, 1990; Anderson, 1997). The “contrast polarity” model is based on the extension of Metelli’s polarity constraint to the more general \textit{x-junction} configurations (see Figure 1), so-called because they are areas where four surfaces meet. Figure 1 shows three gray squares with two overlapping regions, both of which can be transparent or opaque depending on their lightness values and contrasts. Beck and Ivry (1988) noticed that tracing a continuous line following the order of increasing or decreasing lightness values in an \textit{x-junction}, the possible resulting figures always have to be crisscross-shaped, C-shaped, or Z-shaped (see Figure 1).

Adelson and Anandan (1990) and Anderson (1997) point out that these three patterns are crucial for the perception of transparency. While crisscross configurations do not allow the perception of transparency, C-shaped configurations, showing single reversing junctions, lead to unique transparency. In Figure 1B, we observe that the lowermost square is transparent and in front of the uppermost one. With Z-shaped configurations, or non-reversing junctions (Figure 1C), both layers are likely to be perceived as transparent. Several studies have explicitly stated that the perceived depth order of the two overlapping regions should be perceptually ambiguous in this case. Adelson and Anandan (1990) stated that “non-reversing junctions leave the depth-order ambiguous”.

![Figure 1](https://jov.arvojournals.org/pdfaccess.ashx?url=/data/journals/jov/932794/) Figure 1. \textit{X-junctions} are areas where four surfaces meet. The three top panels show \textit{x-junctions} with different contrast polarity arrangements. The three bottom circles show enlarged segments containing the \textit{x-junctions} with a white line, proceeding from the brighter to the darker regions. The leftmost one traces a crisscross shape. Neither the horizontal nor the vertical pair of aligned contour segments preserve contrast polarity. The one in the middle is C-shaped. In this kind of configuration, only one set of aligned contours complies with the contrast polarity model (which in this figure is the horizontal one). The rightmost circle is Z-shaped. In this case, both vertical and horizontal aligned contours comply with the contrast polarity model (adapted from Fleming & Anderson, 2003).
Anderson (1997), consistently with this model, referring to non-reversing x-junctions stated that "which surface appears transparent depends on the perceived depth order, which is also ambiguous". Both Adelson and Anandan and Anderson refer to ambiguity as the property of a stimulus that can possibly be perceived in two alternative ways. This may give rise to a bistable percept of transparency. At any time, a perceptually transparent layer will be in front of the other; but which one it is will differ (Adelson & Anandan, 1990; Anderson, 1997; Kitaoka, 2005). However, these studies do not make specific predictions about the likelihood of occurrence of the two alternative interpretations. In this study, we focus on empirically investigating the observed depth order in non-reversing contrast polarity transparency.

Our experiments will test, firstly, whether non-reversing Z-configurations are sufficient for the perceptual emergence of different depth ordering solutions as predicted by the contrast polarity model. In other words, to what extent does non-reversing contrast polarity lead to ambiguity in depth ordering of transparent layers? To investigate this issue, in all our experiments the stimuli have equal contrast polarity, but vary in lightness. The task is to determine which of two surfaces will be in front of the other. According to the Z-ambiguity prediction, both surfaces are equally qualified for being perceived as “in front” of each other.

Z-configurations alone may not be sufficient for ambiguity. Let us, therefore, consider an additional photometric factor that might determine which surface is “in front”: Anderson’s (1999, 2003) transmittance anchoring principle (TAP). The TAP claims that the visual system treats the highest contrasts regions along their internal or external contours as “transmittance anchors”. This means that they are in plain view, whereas regions with lower values of contrast are characterized by transparency. For example, the x-junction in Figure 2A is C-shaped, meaning contrast polarity violation in the horizontal direction. This leads to unique transparency. According to the anchoring principle, the regions a–b have the highest contrast contours. Therefore, these regions must be in plain view, so the disk p–q must be a transparent object on top of a rectangular surface.

Anderson does not specify what happens if the two regions in a Z-configuration have equal contrast value. It would, however, be plausible to assume there would be approximately equal preference for which surface is “in front”, occluding the other. Let us call these configurations A-configurations. So, besides the “weak” requirement of Z-ambiguity, we may introduce the strong requirement of A-ambiguity as a condition for perceptual ambiguity.

In our Experiment 1, we presented observers with Z-configurations, a subset of which were A-configurations. We tested whether the first (Z-configuration) or the second (A-configuration) condition was sufficient for ambiguity.

In case neither Z- nor A-configurations are sufficient for ambiguity, we may consider a third photometric factor. In Figure 2B, p–q is the region containing the highest contrast contour, and thus, according to the anchoring principle, this region is in plain view. Anderson claimed that the display is likely to be perceived as a transparent surface a–b with a hole p–q cut out of its center (full-layer transparency). However, as Kitaoka (2005) and Koenderink, van Doorn, Pont, and Richards (2008) have shown, full-layer transparency interpretations are quite rare.

In fact, observers might prefer to interpret the display as a half-occlusion, in which a, b, and p are all in plain view, while q is behind the transparent rectangle b. A half-occluded disk might give rise to inconsistent answers with respect to the question whether p–q is in front or behind the rectangle b. Therefore, providing a static display of Figure 2B while asking the same question would give rise to inconsistent answers, even though there may not be any perceptual ambiguity involved.

Consider, by contrast, an animated display in which a uniformly p-colored disk moves across surface a from left to right. Where it hits b, its color turns into q. Observers will be likely to perceive the disk unanimously as slipping underneath surface b, as it continues moving to the right. In general, we may expect inconsistent answers to become more consistent as a result of such animation. Masin (2006), Experiment 2, had shown that motion has no influence on transparency judgments. If so, the increased consistency will not affect a depth preference on average. To test these hypotheses, Experiment 1 included static as well as dynamic conditions.
Finally, when all previous photometric predictions of ambiguity fail, it is important to identify the reason. Experiments 2 and 3 will explore why this is the case for some of our stimuli.

**Experiment 1**

**Methods**

**Participants**

Eleven observers (4 females) with normal or corrected-to-normal vision volunteered to participate in the experiment. All participants were naive to the purposes of the experiment. The ethical committee of the RIKEN BSI had approved the experiment.

**Materials**

It is still a matter of debate what is the critical variable for perceived transparency (Anderson, Singh, & O’Vari, 2008). Until recently, different systems of reference for setting the contrasts between figures were used in perceptual transparency research: reflectance (Metelli, 1985), luminance (Masin, 2006), Michelson contrast (Singh & Anderson, 2002), and lightness values (Beck & Ivry, 1988; Beck, Prazdy, & Ivry, 1984; Kitaoka, 2005). We chose to compute the contrast in terms of lightness values, which allowed us to provide perceptually scaled values of contrast regardless of the luminance levels in our stimuli. We used the following values of lightness: 3.0, 4.0, 4.5, 5.0, 6.0, and 7.0, which correspond to the following average luminance values: 6.97, 12.86, 17.00, 21.14, 26.70, 32.33, and 45.23 cd/m² as measured through a photometer directly pointed to the figures on the screen. This correspondence fits the following conversion equation:

\[ W = 25Y^{1/3} - 17 \]  

where \( W \) is the lightness value, and \( Y \) is the luminance level (Wyszecki, 1963).

Considering that the amount of contrast between regions varies according to the system of reference, the results will include a comparison between lightness, luminance, and Michelson contrast values. We will also calculate the best fit to the corresponding psychometric function for each system.

**Dynamic stimuli**

Sixteen displays of 21.8° × 16.7° visual angle were used, in which a homogeneous achromatic disk of a diameter subtending a visual angle of 1.72° moves from the left to the right side of the display with a constant speed of 4.36° of visual angle per second. In the middle of the display, there is a static rectangle subtending a visual angle of 7.41° horizontally and 8.53° vertically. The disk has two distinct lightness values, one for when it overlaps with the rectangle, and one for when it does not overlap with it. The disk changes its lightness when it overlaps with the left edge of the rectangle, to recover its original lightness when it crosses the right edge of the rectangle. Otherwise, all surfaces keep their original lightness values. The sequence of events is recapitulated in Figure 3, in which the disk is shown at different moments, while partially, totally, and not overlapping with the rectangle.

The lightness of the rectangle, of the disk, and of the disk–rectangle overlapping varied according to experimental conditions, showing one of the following lightness values (luminance values): 3 (6.97 cd/m²), 4 (12.86 cd/m²), 4.5 (17.00 cd/m²), 5.5 (26.70 cd/m²), 6 (32.33 cd/m²), 7 (45.23 cd/m²). The uniform gray background always showed a lightness value of 5 (21.14 cd/m²).

The features of the lightness patterns are given as follows (see also Table 1):

1. Background lightness \( B \) is kept constant and is always in between those of \( R \) and \( D \). Therefore, if \( B < R \), then \( B < D \), or vice versa, if \( B < R \), then \( B > D \).
2. \( O \) is unequal to \( B \) (\( O \neq B \)) and the difference between the contrasts \( RO \) and \( DO \) is always different from zero.
3. \( R \) and \( D \) always have the extreme lightness values in the patterns so either \( R \) is the brightest region and \( D \) is the darkest or vice versa.
4. The highest contrast region is defined as the region showing the greater difference in lightness (or the greater contrast) with all its flanking regions. When

![Figure 3](https://jov.arvojournals.org/pdfaccess.ashx?url=data/journals/jov/932794/)
(\(|R - D| + |R - O| + |R - B|\) > (\(|D - B| + |D - R| + |D - O|\)), the rectangle \(R\) is considered as the highest contrast region, vice versa when (\(|D - B| + |D - R| + |D - O|\)) > (\(|R - D| + |R - O| + |R - B|\)) the disk \(D\) is considered the highest contrast region, finally when (\(|R - D| + |R - O| + |R - B|\) = (\(|D - B| + |D - R| + |D - O|\)) the disk and the rectangle are considered as equally contrasted.

All 16 patterns are Z-configurations, i.e., contrast polarity is preserved along both contours so, according to the contrast polarity model, it should be always possible to interpret their depth ordering in at least the following two ways: either as a disk in front of a rectangle or vice versa. In six patterns, the highest contrast region is the rectangle (\(R\)). Consequently, according to the TAP, the disk should be perceived as transparent and “in front”. In four cases, the rectangle and the disk have the same amount of contrast along surfaces and contours. In other words, they are A-configurations. We should expect ambiguous transparency judgments to be reserved for A-configurations, in case sameness of contrast is used as additional criterion for ambiguity.

### Static stimuli

A variation of the above described dynamic stimuli was used for the static condition. The static stimuli corresponded in region size and lightness patterns to those in the dynamic conditions. The disk was centered in the middle of the screen, with 50% of its entire area overlapping with the rectangle on its right-hand side, leaving the background on its left.

#### Procedure

Participants were seated in a dimly lighted room at a distance of approximately 100 cm from a cathode ray tube monitor (Sony CPD-G520, 21”). In the dynamic condition, after a 500-ms fixation delay, the disk appeared on the left-hand side of the screen, it moved with a constant speed toward the right side of the screen, and then it disappeared on the right-hand side. The entire presentation lasted 5 s and the period in which the disk and the rectangle overlapped lasted 1,666 ms. The moments in which the partial overlapping of the disk and the rectangle shows all the \(B - R - D - O\) lightness values simultaneously in an x-junction-like configuration, last 770 ms in total, 384 ms when \(D\) reaches \(R\) turning into \(O\), and 384 ms when \(O\) leaves \(R\) turning into \(D\) again. After presentation, the instructions appeared on the screen asking participants to respond whether the disk was in front or behind the rectangle. Participants responded by pressing the “\(N\)” and the “\(M\)” keys of a computer keyboard to indicate “behind” and “in front”, respectively. Each lightness condition was repeated ten times, yielding a total of 160 trials presented in random order in a single dynamic condition block.

In the static condition trials, a fixation cross appeared for 500 ms in the center of the screen. Next, the target configuration was presented for 770 ms, a period identical to the combined periods in which an x-junction was visible in dynamic trials. Immediately afterward, a brief tone signal was presented and at the same time the response instruction appeared on the screen. Each lightness condition was repeated ten times, yielding a total of 160 trials. These were presented in random order in a single static condition block. Static and dynamic condition blocks were presented in counterbalanced order with a

### Table 1

<table>
<thead>
<tr>
<th>B–R–D–O lightness pattern</th>
<th>Highest contrast region</th>
<th>Brightest region</th>
<th>Relative lightness</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.0–6.0–3.0–5.5</td>
<td>Disk</td>
<td>Rectangle</td>
<td>DOBR</td>
</tr>
<tr>
<td>5.0–4.0–7.0–4.5</td>
<td>Disk</td>
<td>Disk</td>
<td>ROBD</td>
</tr>
<tr>
<td>5.0–3.0–7.0–4.0</td>
<td>Disk</td>
<td>Disk</td>
<td>ROBD</td>
</tr>
<tr>
<td>5.0–7.0–3.0–6.0</td>
<td>Disk</td>
<td>Rectangle</td>
<td>DBOR</td>
</tr>
<tr>
<td>5.0–5.5–3.0–4.5</td>
<td>Disk</td>
<td>Rectangle</td>
<td>DOBR</td>
</tr>
<tr>
<td>5.0–4.5–7.0–5.5</td>
<td>Disk</td>
<td>Disk</td>
<td>ROBD</td>
</tr>
<tr>
<td>5.0–4.0–5.5–4.5</td>
<td>Equivalent contrast</td>
<td>Disk</td>
<td>ROBD</td>
</tr>
<tr>
<td>5.0–6.0–4.5–5.5</td>
<td>Equivalent contrast</td>
<td>Rectangle</td>
<td>DBOR</td>
</tr>
<tr>
<td>5.0–5.5–4.0–4.5</td>
<td>Equivalent contrast</td>
<td>Rectangle</td>
<td>DOBR</td>
</tr>
<tr>
<td>5.0–4.5–6.0–5.5</td>
<td>Equivalent contrast</td>
<td>Disk</td>
<td>ROBD</td>
</tr>
<tr>
<td>5.0–3.0–5.5–4.5</td>
<td>Rectangle</td>
<td>Disk</td>
<td>ROBD</td>
</tr>
<tr>
<td>5.0–7.0–3.0–4.0</td>
<td>Rectangle</td>
<td>Rectangle</td>
<td>DOBR</td>
</tr>
<tr>
<td>5.0–3.0–7.0–6.0</td>
<td>Rectangle</td>
<td>Disk</td>
<td>RBOD</td>
</tr>
<tr>
<td>5.0–7.0–4.5–5.5</td>
<td>Rectangle</td>
<td>Rectangle</td>
<td>DOBR</td>
</tr>
<tr>
<td>5.0–3.0–6.0–5.5</td>
<td>Rectangle</td>
<td>Disk</td>
<td>ROBD</td>
</tr>
<tr>
<td>5.0–7.0–4.0–4.5</td>
<td>Rectangle</td>
<td>Rectangle</td>
<td>DOBR</td>
</tr>
</tbody>
</table>

The left–right letter position in the rightmost column indicates the relative lightness of the regions so that the elements on the left are darker than those on the right.
break between blocks. The entire experiment lasted approximately 35 min.

Results

Individual responses to each of the lightness patterns varied from 0 (disk never judged as “in front”) to 10 (always “in front”). A 2-factor within-subjects ANOVA with Pattern (16) and Motion (dynamic vs. static displays) was performed. Greenhouse–Geisser correction for sphericity was applied when necessary. Hereinafter, the reported significance values are based on the correction; the degrees of freedom are the original, uncorrected ones. Depth ordering judgments were sensitive to the Pattern,\( F(15, 135) = 93.5, \epsilon = 0.23, p = 0.000, \) but not to Motion,\( F(1, 9) = 0.914, p = 0.36. \) No Pattern \( \times \) Motion interaction was observed,\( F(15, 135) = 0.83, p = 0.64. \)

An additional within-subject ANOVA analysis with disk Lightness (Dark disk, including all configurations in which the disk is the darkest of the displayed figures vs. Bright disk, including all configurations in which the disk is the brightest figure) showed no significant effect of Lightness,\( F(1, 9) = 0.046, p = 0.834, \) indicating that rather than by the absolute lightness of the regions, the depth interpretation is determined by the relation between the lightness values of the four regions.

As shown in Figure 4, depth judgments vary systematically according to the difference in contrast between the disk and all the other regions and between the rectangle and all the other regions. The highest contrast region is judged to lie behind, while the region with the lowest contrast as being the transparent occluder. In the six lightness combinations where the rectangle is the highest contrast region (the six leftmost patterns in Figure 4), observers showed an unambiguous preference for “disk-behind” interpretations. In fact, in these conditions of contrast the perception of “disk in front” was quite low (3% of trials on average). By contrast, in the six rightmost lightness combinations where the disk is the highest contrast region, subjects consistently agreed to interpret the disk as being in front of the rectangle (85% of trials in average). These unbalanced preferences for one or the other depth ordering are at variance with the contrast polarity model.

Interestingly, when the sum of contrasts of \( R \) is equal to the sum of contrasts of \( D \) (as in the central four patterns in Figure 4), depth order perception was ambiguous, with 40% of “disk-in-front” judgments.

A one-factor ANOVA on the standard deviations of the individual mean judgments showed that the static condition led to depth judgments that are more scattered around the mean than the ones in the dynamic condition,\( F(1, 15) = 5.08, p = 0.83. \)

Discussion

According to the contrast polarity model (Adelson & Anandan, 1990), we expected to find that our Z-patterns, in which contrast polarity is always preserved along both contours, cause the emergence of a considerable degree of ambiguity in the perception of the depth order of the rectangle and the disk. The prediction of the contrast polarity model does not fit the data. Most of the patterns,
with a few relevant exceptions, elicited an unambiguous preference for a depth interpretation. This result is consistent between observers. In other words, Z-ambiguity is not a sufficient criterion for perceiving transparent configurations as ambiguous.

According to the TAP (Anderson, 2003), we expected to find that the highest contrast region would be perceived in plain view and consequently to be viewed as continuing behind the semitransparent occluder. Our results confirm this expectation. In the six patterns in which the rectangle had the highest contrast, observers consistently perceived the disk in front of the rectangle. In the six cases in which the disk had the highest contrast with the adjacent regions, observers tended to perceive the rectangle in front of the disk. The remaining four cases are A-configurations, in which the disk and the rectangle were equally contrasting with the adjacent regions. Here, depth judgments show different degrees of preference for one or the other interpretation. This result is consistent with TAP: all A-configurations are perceived as ambiguous to some degree.

The fact that these results occur equally for static and dynamic regions offers a reassurance that half-occlusion in static displays plays no significant role in depth judgment. Previous studies found no effect of motion in perception of transparency (Masin, 2006; Experiment 2). However, we observed a significant lower standard deviation in depth judgments of dynamic than static displays; this indicates that motion reduces the inconsistency in depth ordering responses. In Experiment 2, only dynamic configurations were therefore used.

Despite the occurrence of ambiguity in accordance with TAP, the thesis that A-configurations are sufficient for ambiguity needs modification. There is considerable variation in preference of A-configurations, which TAP leaves unaccounted for. The region that seems to play a prominent role is region O, the overlap of rectangle and circle (see Figure 3). In Figure 4, observers tended to perceive the region with the smaller contrast with O as in front of the other. Specifically the rectangle is likely perceived in front in the B–R–D–O configurations 5.0–4.0–5.5–4.5 and 5.0–4.5–6.0–5.5, while the disk is more often perceived in front in configurations 5.0–6.0–4.5–5.5 and 5.0–5.5–4.0–4.5. This hypothesis is further tested in Experiment 2.

## Experiment 2

Results of Experiment 1 suggested that, for judging depth stratification of overlapping regions with non-reversing transparency, participants use contrast differences. In order to confirm this data, in the second experiment we systematically varied the difference between the sum of contrasts of region O with the adjacent regions C and B.

### Methods

#### Participants

Ten observers (5 females) with normal or corrected-to-normal vision participated in the experiment. All participants were naive to the purposes of the experiment. The ethical committee of the RIKEN BSI had approved the experiment.

#### Materials

Eight different lightness patterns, each consisting of a rectangle and a moving disk with the same figural features of the dynamic stimuli of Experiment 1, were used as stimuli. As in Experiment 1, the various elements of the scene can show one of the following lightness values: 3, 4, 4.5, 5.5, 6, 7. The uniform gray background (B) always has a lightness value of 5. In each display, the lightness of the rectangle (R), of the disk (D), and of the disk–rectangle overlapping (O) varies according to the experimental conditions.

The B–R–D–O lightness patterns varied according to the same general criteria of Experiment 1. As shown in Table 2, we systematically varied the difference of the sum of contrasts of the disk and the rectangle in such a way that the resulting patterns have the difference |RB| + |RO| − |DB| + |DO| varying from −3 to +3 steps in the lightness scale.

#### Procedure

The procedure was the same as in the dynamic condition of Experiment 1. Each condition of lightness was repeated ten times for a total of 80 trials.

### Results

We ran a within-subjects one-factor ANOVA with contrast difference (8 B–R–D–O lightness different
combinations) as within factor. Participants’ depth ordering judgments varied in function of the difference in the total amount of contrast showed by the rectangle and the disk; $F(7, 70) = 77.85$, $\varepsilon = 0.34$, $p = 0.000$.

Post-hoc analyses (Bonferroni tests) showed that the three patterns in which the sum of contrasts of D was larger than the sum of contrasts of R (with a difference of 3, 2, and 1 steps on the lightness scale) all lead to an unambiguous and consistent “disk-behind” depth interpretation (average “disk-in-front” judgment = 0.15). The fact that these three patterns did not differ from one another in their depth interpretation indicates that the magnitude of the contrast difference is irrelevant. All three patterns in which the contrast of R was larger than the contrast of D lead observers to view the disk as passing in front of the rectangle in nearly all cases (average “disk-in-front” judgment = 0.88). Depth interpretations of the two configurations with equivalent R and D contrasts significantly differ from each other ($p < 0.001$), indicating that A-configurations are not sufficient to perceive complete bistability in disk–rectangle depth ordering. This result confirms that of Experiment 1. In A-configurations, the region that has the smaller contrast with the overlapping region O is the one that is perceived “in front”. In particular, in the R–B–D–O configuration “5.0–3.0–6.0–4.0”, with RO smaller than DO, the disk is consistently perceived behind the rectangle as often as in patterns in which the contrast of R was larger than that of D. However, in the configuration 5.0–3.0–6.0–4.0, where DO is smaller than RO, depth ordering perception seems to be more ambiguous (mean judgment = 4.45).

In Figure 5, the figure with the highest contrast always tends to be perceived as in front, irrespectively of the magnitude of the contrast difference. The shape of the plotted data points in Figures 4 and 5 suggests that perceptual depth preferences follow a psychometric function of all the contrast differences in the pattern. We therefore performed for all configurations of Experiments 1 and 2 a non-linear regression analysis. The result was a very good fit of our data with the sigmoid function in

$$y = \frac{a}{1 + e^{-\left(\frac{x-x_0}{b}\right)}},$$

where $a = 8.61$, $b = 0.045$, and $x_0 = 0.02$; Rsqr = 0.9 (Figure 6, bottom-left graph). This result allows us to interpret the depth judgments in terms of a critical choice threshold.

**Metrics for perceived contrast**

When describing our experimental materials, we preferred lightness measures to other possible scales, as lightness values are perceptually scaled and are convenient for choosing contrast values while avoiding problems of saturation. Here we converted the lightness values of all our patterns to luminance and Michelson contrast. Figure 6 shows the depth order preferences plotted as a function of luminance and Michelson contrast values, besides the already mentioned lightness contrasts. In each of the three plots, at the bottom of the figure, we show how the most contrasted figure tends to be judged in front, irrespectively of whether contrast is based on lightness, luminance, or Michelson measures. Non-linear regression analyses indicates a good fit with the sigmoid function shown in Equation 2 for each one of the three plots, with Rsqrs = 0.82 for luminance (with parameters $a = 8.59$, $b = 4.11$).

![Figure 5](https://jov.arvojournals.org/pdfaccess.ashx?url=/data/journals/jov/932794/)

Figure 5. Proportion of “disk-in-front” judgments as a function of $|RB + RO| - |DB + DO|$ contrast difference. Vertical bars indicate standard errors.
and $x^0 = 1.26$), and Rsqr = 0.73 for Michelson contrast (with parameters $a = 9.01$, $b = 0.14$, and $x^0 = 0.04$). Our observation that a choice threshold is involved is therefore independent of the reference system. However, lightness provided the best fit (Rsqr = 0.9), suggesting that lightness contrast is likely to be the critical photometric value that our visual system uses for reconstructing depth order of transparent layers.

**Discussion**

Depth order preference abruptly changes according to the identity of the most contrasted figure, rather than proportionally to the amount of contrast. High degrees of ambiguity, therefore, are found only in those configurations in which the contrast of the two figures is equal. However, in those patterns in which there is no unique
highest contrast region, the layer with smaller contrast with the overlapping region tends to be perceived as in front. All results are most adequately expressed in terms of lightness contrast, but they do not depend on this system.

These results supply the sufficient conditions for ambiguity not provided by Anderson’s (2003) TAP. We might, however, consider an alternative interpretation in which in configurations with non-reversing contrast polarity, the contrasts of two layers with their overlapping region is the only information used to establish their depth order. This way, observers identify the region with the smaller contrast with the overlapping region as transparent and in front of the other. Thus, contrary to TAP, no further contrast between figures and background would be necessary for depth order judgment. However, in this case we would expect no difference in depth order preference between patterns of which B–R–D–O lightness values are, and in front of the other. Thus, contrary to TAP, no further contrast between figures and background would be necessary for depth order judgment. However, in this case we would expect no difference in depth order preference between patterns of which B–R–D–O lightness values are, respectively, 5.0–6.0–3.0–4.0 and 5.0–6.0–4.0–4.5. In fact, in both these patterns the difference RO respectively, 5.0–6.0–3.0–4.0 and 5.0–6.0–4.0–4.5. In all configurations, the difference |RB + RO| − |DB + DO| is equal to zero, so that the overlapping regions are equally contrasted with the adjacent contours and surfaces.

In order to allow a better control of the relative size of the overlapping of the disk and the quadrilateral figure, we replaced the rectangle with a square, which we placed in the middle of the screen. The dimensions of the square are constant in all trials, and its side subtends a visual angle of 3.64°. The disk size changes according to four conditions. In a first “small” condition, the disk has a diameter subtending a visual angle of 0.93°, corresponding to 0.25 of the side of the square. In a second “medium” condition, the diameter subtends a visual angle of 2.3°, corresponding to 0.63 of the side of the square. In the third “same-size” condition, the diameter subtends the same visual angle of the side of the square (3.64°), so that when overlapping with the square, for a moment the disk perfectly overlaps with the square. In this condition, only the four corners of the square remain unchanged in lightness when the two figures completely overlap. In the fourth “exceeding” condition, the diameter of the disk subtends a visual angle of 4.6° (1.26 times the side of the square), so that the disk exceeds the dimensions of the square. Consequently, when the center of the disk and the center of the square overlap, the lightness of the entire square changes.

Experiment 3

In Experiments 1 and 2, participants showed a general tendency to perceive the rectangle as transparent and in front of the disk, even though the sum of contrasts is equivalent. In Experiment 3, we verify the hypothesis that the figural factor of the relative size of the layers can be the cause of this bias.

Methods

Participants

Eleven observers (4 females) with normal or corrected-to-normal vision participated in the experiment. All participants were naive to the purposes of the experiment.

The ethical committee of the RIKEN BSI had approved the experiment.

Materials

Four different lightness pattern displays similar to the ones used in the previous experiments were used. The following B–R–D–O lightness patterns were used: [5.0–4.0–5.5–4.5], [5.0–4.5–6.0–5.5], [5.0–6.0–4.5–5.5], [5.0–5.5–4.0–4.5]. In all configurations, the difference |RB + RO| − |DB + DO| is equal to zero, so that the overlapping regions are equally contrasted with the adjacent contours and surfaces.

Results

We ran a two-factor within-subjects ANOVA with Lightness Contrast (4 levels) and Size (small, medium, large, exceeding) as factors. The dependent variable was the “disk-in-front” proportion of judgments.

Confirming the results of the previous experiments, the main factor Lightness Contrast was significant, F(3, 30) = 52.9, e = 0.85, p = 0.000. As shown in Figure 7, Bonferroni post-hoc test confirmed that the region exhibiting a lower contrast with the overlapping region
O is significantly more frequently judged to be “in front” than the region showing a higher contrast with O.

The main factor Size was also significant, \( F(3, 30) = 18.66, \ v = 0.65, \ p = 0.000 \) (see Figure 7). The size of the disk clearly influences depth order judgment. Post-hoc analysis (Bonferroni test) showed that the displays in which the size of the disk exceeds the size of the rectangle are perceived more frequently as “in front” than all the other displays, which do not differ from one another. The contrast-vs.-size interaction is also significant, \( F(9, 90) = 4.26, \ v = 0.47, \ p = 0.000 \). Post-hoc comparisons indicate that the size of the disk does not affect all the lightness patterns in the same way. In particular, in the two conditions of lightness in which DO \( \leq \) RO, the medium-sized disk is more frequently perceived “behind” than the smallest disks. For example, in configuration 5.0–4.5–6.0–5.5 the large disk is perceived more often “in front” than the medium-sized disk, but not than the smallest disk. A more linear trend is shown for the configurations in which RO \( < \) DO, where the size of the disk causes a significant shift toward the “disk-in-front” judgments only in the case of disks exceeding the size of the rectangle.

**Discussion**

The visual system, besides photometric information, also uses non-photometric information to dissect the transparency (see Fleming & Anderson, 2003 for a review). Previous studies found the importance of other non-photometric factors such as good continuation (Beck & Ivry, 1988; Metelli, 1974) and subjective contours (Nakayama, Shimojo, & Ramachandran, 1990) in transparency and depth order judgment. In Experiment 3, we showed that the relative size of surfaces is another factor to be taken into account. For given lightness conditions, the smaller figure, which completely changes its lightness value during the overlapping phase, tends to be judged in front of the other region. This result can explain why in the previous experiments, where the disk was always smaller than the rectangle, a systematic bias was obtained toward a “disk-behind” interpretation. Further experiments with different lightness, figural shape, and dynamicity conditions should be conducted to reach a better understanding of the role of the relative size of figures and its interactions (and possible conflicts) with other non-photometric and photometric factors in depth ordering of transparent layers.

**General discussion**

We studied perceptual transparency and depth ordering by experimentally testing some of the principal assumptions of two main models on the topic, namely Adelson and Anandan’s (1990) contrast polarity model and Anderson’s (1999, 2003) Transmittance Anchoring Principle (TAP).

The contrast polarity model claims that x-junction configurations with non-reversing contrast polarity (Z-configurations) generate ambiguity in depth order judgments. This turned out to be the case only for a
subset of patterns close to a critical threshold; the others elicit clear preferences for a certain depth ordering.

These preferences were found to be in accordance with TAP and consistent with the extensive experimental evidence for this principle (Anderson, 1999; Kitaoka, 2005; Oyama & Nakahara, 1960). In our Experiments 1 and 2, the highest contrast layer is almost always interpreted as the region in plain view. However, not all our Z-configurations have a unique highest contrast layer. Some were chosen to be, in addition, A-configurations, i.e., configurations in which the contrast of the two layers is equivalent. TAP did not contemplate such cases. They show considerably more ambiguity, but still a preference is accorded to one of the interpretations. Evidently, observers must rely on additional criteria besides reversing x-junctions or anchoring to determine their percept.

As an additional photometric criterion for A-configurations, we proposed the amount of contrast between the two layers and their overlapping region. The surface showing the lower contrast with the region of overlapping tends to be perceived as transparent and, consequently, in front. This principle was shown to operate in the narrow region around the critical decision threshold where ambiguity occurs.

Our experiments show that neither non-reversing contrast polarity nor anchoring are sufficient conditions for observers to consider two possible perceptual transparency interpretations. Perceived transparency seems more stable than what these models theoretically predict.

We were able to fit this result with a psychometric function and identify the critical response threshold. This result was unaffected by the precise coordinate system used to determine contrast. Lightness, luminance, and Michelson contrasts all yielded good fits to a psychometric curve. No matter what coordinate system is used, therefore, observers are only confined to ambiguity in the narrow region where two regions have equal contrast with their overlap and background. The great consistency in judgment as a result of photometric factors illustrates the fundamental role of the mechanisms involved in resolving transparency for visual information processing.

In addition to photometric factors, transparency judgments are influenced also by non-photometric factors such as motion. In Experiment 1, it was shown that motion does not change the average preference for any interpretation, but it increases the consistency of depth judgments across trials. Motion, therefore, is a factor that seems only to contribute to the consistency of transparency judgments.

The same could not be said for the other non-photometric factor we investigated: size. When one surface is much smaller than another, such that the latter completely envelops it, its lightness information is not directly available to the observer. Experiment 3 shows that in ambiguous lightness conditions, such a surface is more likely to be perceived as behind. These effects are not restricted to ambiguous lightness conditions; in fact in Experiments 1 and 2, the size effect seems to explain why the disk was not always judged as in front, where lightness conditions would have predicted that.

Conclusion

We investigated the role of photometric factors in depth ordering judgment of displays characterized by transparency. We found that neither contrast polarity model nor TAP can fully account for how the visual system organizes the depth ordering of transparent layers. In most of the cases where preference is left indeterminate by both these principles, the visual system makes precise and coherent decisions about depth ordering of multiple planes in the presence of transparency. There is little room left for depth ordering ambiguity. We observed a critical choice threshold, in which the highest contrast surface is perceived to be transparent and in front. Ambiguity is observed only around this critical threshold. We introduced an additional photometric criterion, which is based on the contrast between a surface and the region where it overlaps with another. This contrast is able to modulate depth preference in the ambiguous range. In this range, non-photometric factors such as motion and size could also play a predominant role. These observations are independent of the precise coordinate system used to determine the contrast.

Acknowledgments

The authors thank Micheal Landy and the anonymous reviewers for helpful comments on earlier drafts of the manuscript.

Commercial relationships: none.
Corresponding author: Dr. Franco Delogu.
Email: francedelogu@brain.riken.jp.
Address: 2-1 Hirosawa, Wako, Saitama, Japan.

References


