# Gestalt and phenomenal transparency

**Jan Koenderink,[1,*] Andrea van Doorn,[1] Sylvia Pont,[1] and Whitman Richards[2]**

[1]*Universiteit Utrecht, Physics and Astronomy, Princetonplein 5, 3584CC Utrecht, The Netherlands*
[2]*Massachusetts Institute of Technology Building 32-364, 77 Massachusetts Avenue, Cambridge, Massachusetts 02139, USA*
**Corresponding author: j.j.koenderink@phys.uu.nl*

Received March 13, 2007; revised July 10, 2007; accepted August 20, 2007; posted October 31, 2007 (Doc. ID 81017); published December 19, 2007

Phenomenal transparency is commonly studied by using a stimulus configuration introduced by Metelli: a bipartite patch, divided into equal left and right halves is overlaid with a smaller, concentric bipartite patch, divided along the same line. Observers are instructed to report either a transparent patch over an opaque bipartite field or a mosaic of four opaque patches. We show theoretically and empirically that these are only two of five generic perceptual categories, namely, transparent patch, transparent annulus (hole), mosaic, partial transparency, and multiple transparency (ambiguous) cases. Thus Gestalt factors complicate the interpretation "phenomenal transparency." We propose a framework that avoids this complication. There is excellent agreement between predictions and results. © 2007 Optical Society of America

*OCIS codes:* 330.0330, 330.5510, 350.2450, 290.7050.

## 1. INTRODUCTION

Phenomenal transparency has been used for striking demonstrations where one apparently sees through opaque objects. It sometimes occurs in natural situations, even on the scale of the landscape, where it gives rise to the perception of "glass mountains" (Metzger [1,2]). The studies of Metelli [3–6] have become classics in the field. Here we study a particular configuration that has often appeared in Metelli's and later research.

Although generally known as "phenomenal transparency," the transparent layer actually has to be understood as translucent; i.e., it apparently also scatters radiation to the eye. Thus by phenomenal transparency we refer to phenomenal translucency in this paper. Notice that a transparent sheet, illuminated from the viewing direction, can only darken what is behind it, whereas a translucent sheet can also lighten what is behind it.

The Metelli figure (Fig. 1) consists of four colored areas that abut at four contours where two areas meet and two vertices where four areas come together. In this paper we only consider monochrome patterns.

The union of the areas is a square. The areas are divided into two types, inner and outer. Since the geometrical configuration is bilaterally symmetric about the vertical midline, one distinguishes left inner, right inner, left outer, and right outer patches. The union of the inner patches is a small square, concentric with the large square (that is the union of all areas). The union of the outer areas is a (square) annulus. Depending on the intensities of the four regions, the human observer perceives different phenomenal configurations. One possibility is evidently a mosaic of four opaque abutting patches (like a jigsaw puzzle). Most of the literature considers another possibility in which the large square appears as two abutting opaque rectangles, overlaid with the small square, which appears as a single patch. The patch is seen as undivided, even though the two inner areas have different intensities. The difference is accounted for by seeing the small square as translucent; thus the bipartite background shines through, leading to an intensity difference that does not require the small square to consist of two distinct parts. Notice that the perception involves the attribution of both a depth order and a material property.

It is often suggested [3] that there exist only these two distinct perceptions when viewing this figure, the particular perception obtained being a function of the lightnesses of the four uniform areas. Some accounts in the literature hint at alternative perceptions involving partial translucency [7–9]. In at least one study [10] a larger number of categorically different perceptions was assumed.

Metelli proposed a simplified model of the physics of turbid layers that might account for the perceptions of human observers [4,11,12]. The theory involved in his predictions applies this simplified physics to *a priori* assumed configurations of layers.

A more general approach [10,13–15] would be to attempt to list all possible geometrical interpretations of the scene and apply either Metelli's model of the optics or another, perhaps more ecologically valid, model (e.g., Kubelka–Munk analysis [16]) in order to check these interpretations for their physical possibility. Such an approach is difficult because the number of possible interpretations is very large and possibly involves additional scene parameters. (For example, is the stack of layers frontally illuminated, or is there a backlight in addition?) Moreover, it has to be expected that multiple interpretations might apply; that is to say, a theory of greater generality is likely to predict ambiguous perceptions.

In this paper we attempt such a more general analysis. Because the predictions lead to a larger number of percep-

 

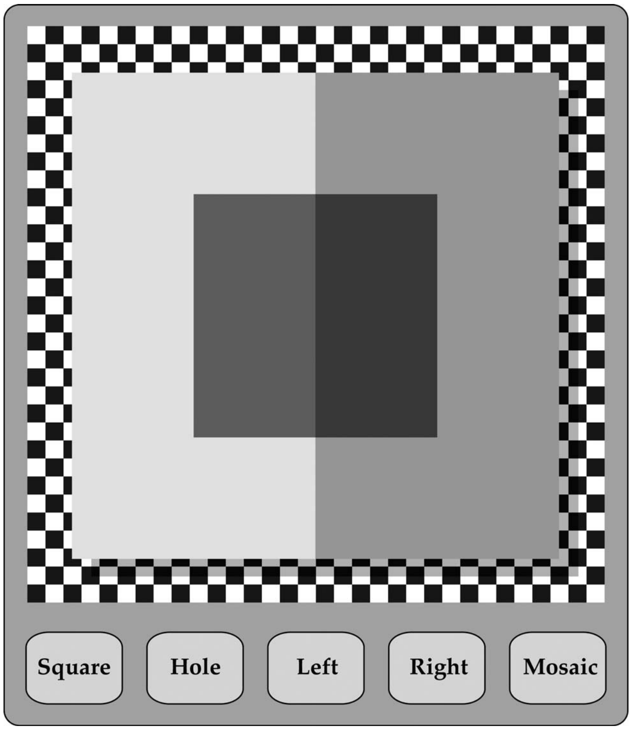


Fig. 1. Example of a stimulus as used in the experiment. The background serves to anchor the basic Metelli figure. The observer selects a category by clicking the corresponding button with the mouse. Immediately after the selection the next stimulus appears.

tual categories than is conventionally considered, we also present novel psychophysical material taking such complications into account.

## 2. THEORETICAL ANALYSIS OF THE METELLI FIGURE

If possible amodal completions are included, the number of layered representations in terms of depth stacks of opaque sheets that might be seen is very large. Many such representations are never seen, though. The laws of Gestalt [1,2,17–21] forbid them.

The number of possible interpretations increases if the layers are permitted to be transparent or translucent. The analysis then requires that the laws of Gestalt be augmented with rules determining the possibilities for such translucency as originally attempted by Metelli [4]. Metelli was much interested in Gestalt phenomena and has left an impressive body of work on the topic [9,22–31]. However, our aim is different in that we merely attempt to outline a minimum set of categories that will allow observers of the Metelli configuration to categorize a particular stimulus without ever feeling forced to make an unlikely choice.

### A. Geometrical Structure

We consider two basic principles that will allow us to select the viable representations of the Metelli figure. They are [32]

**Occam's Razor:** The representation should not assume structural elements unless a reason exists;

**Genericity**: The representation should be structurally stable, i.e., not contain accidents of vanishing probability [33].

Since occlusion may be assumed at no cost, the rules may generate complicated shapes. Occam's razor implies that the simpler shapes postulated under an interpretation of partial occlusion are to be preferred over cut out shapes. The simplest shapes in the Metelli figure are squares and rectangles in canonical orientation, these may be considered to be roughly equally simple. The nonconvex C-shaped or even more complicated (by union of disjunct parts) shapes that arise in the analysis are reckoned more complex than convex shapes such as rectangles. Translucency, too, may generate complex shapes from simple ones and works just as occlusion does in that respect. Occlusion or translucency must respect genericity; i.e., a boundary that runs below a translucent sheet must show good continuation, and so forth.

Thus the viable representations of Metelli's figure are composed of physically homogeneous sheets (either opaque or translucent) in rectangular or square shapes, stacked in depth, in a general position (the technical term is "generic configuration," e.g., see [34]). This restricts the possible interpretations to only a few that can be easily enumerated.

One representation that is always possible is that of a mosaic of four abutting and perfectly fitting pieces. This will be the default representation if no simpler one can be found. If simpler representations exist, the mosaic representation will have to be discarded in favor of the simplest one. Notice that it may be the case that $n>1$ equally simple representations exist. In such a case all $n$ of these will be retained and the resulting representation declared to be ambiguous with $n$-fold degeneracy.

The geometrical analysis can be done *a priori*. Whether the geometrical representations are actually viable depends on whether the translucency relations that have to be assumed are actually possible, e.g., whether they satisfy the laws of physics for frontally illuminated turbid layers.

It is easily verified that the above principles allow five distinct geometrical representations (including the default one).

- A translucent small square in front of a bipartite structure composed of two abutting rectangles (Fig. 2, left). (The bipartition is to be understood as an opaque rectangle in front of an opaque square, but the precise structure need not be further elucidated.)

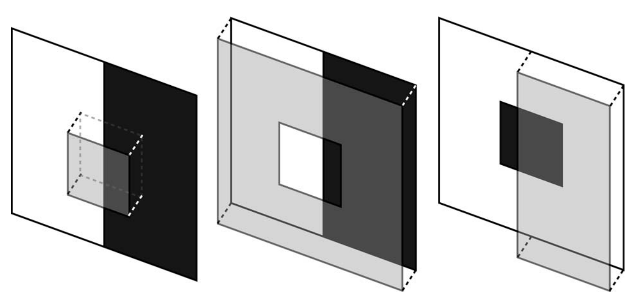

Fig. 2. Generic interpretations of the Metelli figure involving a translucent layer.

- A translucent sheet with small square hole in front of a bipartite structure composed of two abutting rectangles (Fig. 2, middle). (The bipartition is to be understood as an opaque rectangle in front of an opaque square, but the precise structure need not be further elucidated.)
- A translucent rectangle at the right-hand side in front of a bipartite structure composed of a large square annulus and a concentric small square (Fig. 2, right). (The bipartition is to be understood as an opaque square in front of an opaque square, but the precise structure need not be further elucidated.)
- A translucent rectangle at the left-hand side in front of a bipartite structure composed of a large square annulus and a concentric small square (mirror image of Fig. 2, right). (The bipartition is to be understood as an opaque square in front of an opaque square, but the precise structure need not be further elucidated.)
- The default mosaic structure. (The mosaic is to be understood as a stack of opaque squares and rectangles, but the precise structure need not be further elucidated.)

Notice that the representations are only partially determined. For instance, a bipartite, opaque background could consist of two opaque abutting rectangles or an opaque square overlaid with an opaque rectangle on one side. However, such remaining ambiguities are irrelevant in the present context. The relevant aspect is the frontal, translucent sheet, which can be a central square or hole, a left or a right rectangle, or may not be present.

### B. Optical Structure

The physics of a thin turbid layer in front of an opaque, Lambertian surface, diffusely illuminated from the front, is approximately described by Kubelka–Munk two-flux theory [16]. It is assumed that the layer has no additional Fresnel reflection, that is to say, that it has the refractive index of air. A sheet of blotting paper (clear cellulose fibers randomly suspended in air, like a felt) comes close as a physical implementation. The incident flux of radiation propagates through the sheet, suffering absorption and multiple scattering as it progresses. When it emerges from the turbid medium it is partially scattered back into that medium by the backing, opaque surface. The flux that emerges from the turbid medium and enters the eye is made up of two components.

- Radiation that was scattered backwards and emerges from the medium. It has suffered some absorption and scattering in the turbid medium. When the turbid layer is (optically) infinitely thick, this is all one gets.
- Radiation that emerged from the turbid layer, was scattered back by the backing substrate, and made it all the way (backwards) through the turbid medium. This contribution occurs only for (optically) thin layers. It evidently carries an imprint from the backing substrate.

Notice that the radiation is randomly scattered in the turbid medium and may have been scattered by the substrate on multiple occasions.

In Metelli's approximation [5] the two components are simply linearly combined with coefficients that parameterize this phenomenological description. Thus you write for the diffuse reflectance $r$ of the turbid layer on a substrate of reflectance $t$,

$$r = (1-\alpha)\varrho + \alpha t, \tag{1}$$

where $r_0=(1-\alpha)\varrho=\sigma$ is the reflectance of the turbid layer on a black substrate ($t=0$). The expression $(1-\alpha)\varrho$ instead of simply $\sigma$ is a convention that is ultimately derived from Metelli's laboratory apparatus, the episcopister. It is conceptually perhaps somewhat simpler to regard $\sigma$ and $\alpha$ as parameters in a linear approximation of the physical optics.

The parameter $\sigma$ vanishes for a layer that does not scatter; it depends on the scattering power of the material, its absorption, and the thickness of the layer. It may be called the turbidity.

The reflection on a white substrate is $r_1=\sigma+\alpha$. Notice that the contrast

$$H = \frac{r_1 - r_0}{r_1 + r_0} = \frac{\alpha}{2\sigma + \alpha} \tag{2}$$

measures the extent to which a black–white contrast is hidden by the turbid layer; it may be called the "hiding power." The turbidity and the hiding power are phenomenological parameters that describe the material properties (in this case optical) of the sheet.

Because the reflectance is nonnegative and is equal to or less than unity, this expression is subject to the constraints

$$0 \leqslant r, r_0, r_1, \sigma, \alpha \leqslant 1. \tag{3}$$

Suppose you have a bipartite background, with reflectances $x$, $y$ (say). A turbid layer, characterized with turbidity $\sigma$ and hiding power determined by $\alpha$ is superimposed, revealing the substrate with reflectance $x$ as a reflectance $u$ and the substrate with reflectance $y$ as a reflectance $v$ (say). Then

$$u = \sigma + \alpha x, \qquad v = \sigma + \alpha y, \tag{4}$$

from which you find that

$$\alpha = \frac{v-u}{y-x}, \qquad \sigma = \frac{uy - vx}{y-x}. \tag{5}$$

Thus the configuration $\{\{x,y\},\{u,v\}\}$ is a possible one if

$$0 \leqslant \frac{uy - vx}{y-x} \leqslant 1, \qquad 0 \leqslant \frac{uy-vx}{y-x} + \frac{v-u}{y-x} \leqslant 1. \tag{6}$$

These constraints may be simplified and written in a variety of convenient formats. We prefer the form (using the conventional logical connectives $\wedge$ for AND and $\vee$ for OR)

$$((x>y)\wedge(u>v)\wedge P)\vee((x<y)\wedge(u<v)\wedge Q), \tag{7}$$

with

$$P = \left(\left((u<x)\wedge\left(\frac{u}{x}<\frac{y}{v}\right)\right)\vee(x+yu+v>y+u+xv \geqslant x)\right), \tag{8}$$

$$Q = \left( \left( (u < x) \wedge \left( \frac{u}{x} > \frac{v}{y} \right) \right) \vee (u \geqslant x + yu + v < y + u + xv) \right). \quad (9)$$

Notice that constraint (7) is nonlinear: boundary surfaces are either planar or patches of a hyperboloid of one sheet. The constraint bounds a curvilinear polyhedral region in the four dimensional hypercube $0 \leqslant x, y, u, x \leqslant 1$ (Fig. 3).

Notice that the expression for constraint (7) in terms of the two pairs of parameters $\{\{x,y\},\{u,v\}\}$ is valid quite independent of any particular Gestalt interpretation. In the conventional interpretation (i.e., Metelli's and most follow-up literature) one would identify the left-hand outer region with parameter $x$ and the right-hand outer region with the parameter $y$. These would be regarded as opaque background regions. The left-hand inner region would be identified with the parameter $u$, and the right-hand inner region with the parameter $v$, the notion being that these areas represent a single translucent overlay such that $u$ is $x$ as seen through the overlay and $v$ is $y$ as seen through the overlay.

However, there is no pressing reason to allow the interpretation to be limited to this case. The constraint applies to any generic Gestalt interpretation. In each interpretation the constraint yields the condition that that particular Gestalt interpretation is a possibility. The various configurations illustrated in Fig. 2 lead to different assignments of the intensities $x$, $y$, $u$, and $v$ to the four regions (see Fig. 4). For instance, in the conventional case of Metelli's phenomenal transparency, the small square is seen as a translucent overlay over an opaque bipartite large square. The left- and right-hand outer areas are assigned $x$ and $y$, respectively. Then the left-hand inner area is the left-hand outer area shining through the small square and is thus assigned $u$, whereas the right-hand inner area is the right-hand outer area shining through the small square and is thus assigned $v$ (Fig. 2, left, and Fig. 4, left).

For the hole category one has to interchange $(x,y)$ with $(u,v)$, since now the opaque ground is seen in the inner square, whereas the outer annulus is seen as an undivided translucent sheet with a hole (Fig. 2, center, and Fig. 4, second from the left). Similar reasoning applies to

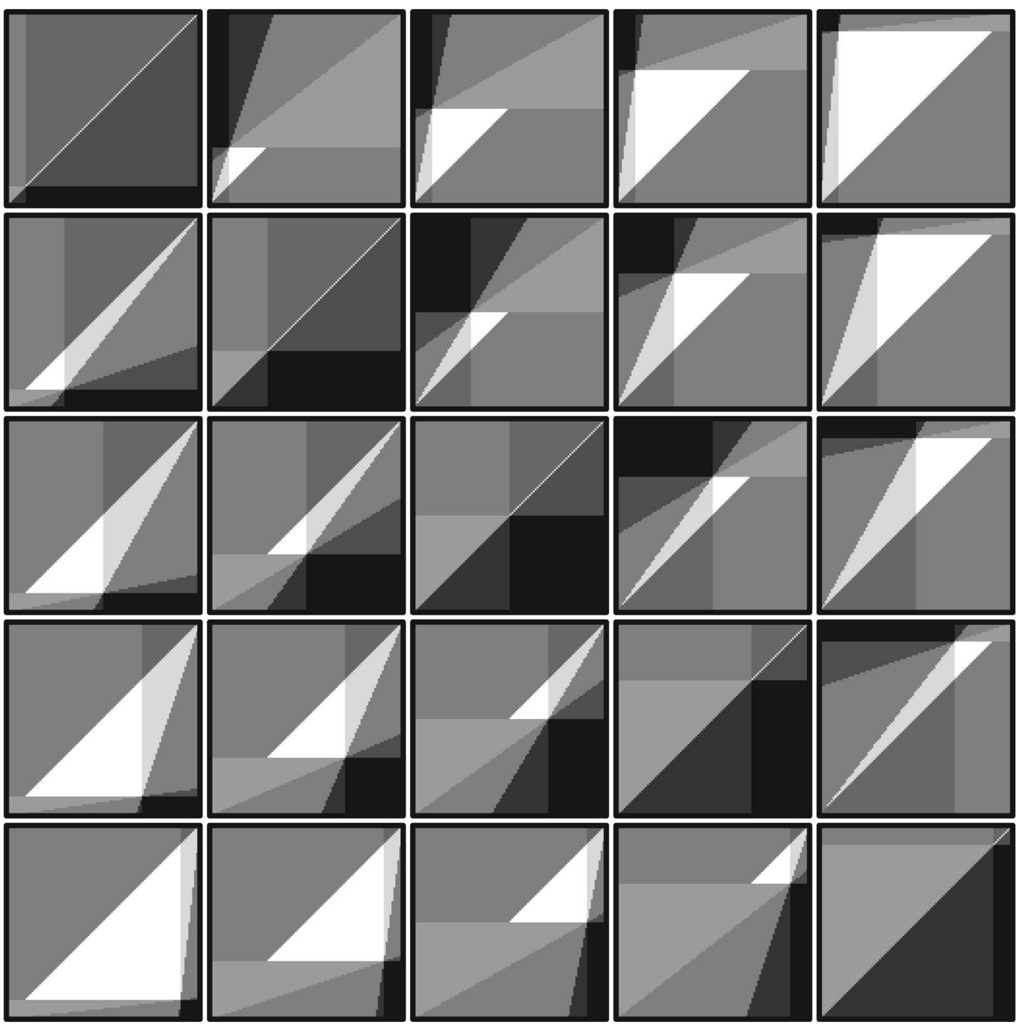

Fig. 3. Cross sections through the *XYUV* hypercube. In the array of cross sections you have $x=0.1, 0.3, 0.5, 0.7, 0.9$ from left to right and $y=0.1, 0.3, 0.5, 0.7, 0.9$ from bottom to top. In each cross section $u$ runs horizontally, and $v$ vertically from zero to one. The gray tones signify (from lightest to darkest) Square (white), Square or Right, Square or Left, Right, Mosaic, Left, Hole, or Right, Hole or Left, and Hole (black).

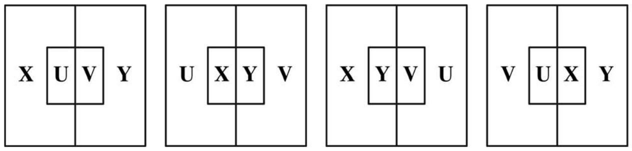


Fig. 4. Designations of X, Y, U and V regions for the Square, Hole, Right and Left translucency case (from left to right).

the left- and right-hand partial transparency cases (Fig. 2, right, and Fig. 4, third and fourth from the left).

As it turns out, the constraint is sometimes satisfied for two distinct Gestalt configurations simultaneously. In such cases either interpretation is a possible one. We may speak of multiple translucency. Since we permit only five response categories in the experiment (excluding cases of multiple translucency), these cases are ambiguous in the sense that two correct responses are possible for a single stimulus.

If the constraint is not satisfied for any of these assignments, the configuration has to be classified as a mosaic; the perception of four, opaque, abutting patches is the simplest explanation. All in all there exist nine distinct Gestalt interpretations, of which four are cases of multiple translucency, two cases of partial translucency, two of translucency, and one of no translucent interpretation. [Remember that the bulk of the literature considers only the transparent square (usually denoted transparent) and the mosaic (usually denoted not transparent) categories.]

We show examples of various cases in Fig. 5.

### C. Generic Interpretations (Gestalts) of the Metelli Figure

For each of the generic geometrical interpretations of the Metelli figure one may verify whether the Metelli constraint has been satisfied. If so, then that interpretation is indeed a possible one; if not, it has to be discarded. If only one interpretation survives this process, then this is the predicted response in a psychophysical experiment. If several interpretations survive the process, then the prediction is that the percept will be an ambiguous (possible multistable) one, and that any of the possible interpretations may be expected. Notice that the mosaic interpretation will always survive, since it involves no translucency at all. However, it is an acceptable interpretation only if the ones involving translucency fail, being complicated and containing accidental features. It is a mere default interpretation.

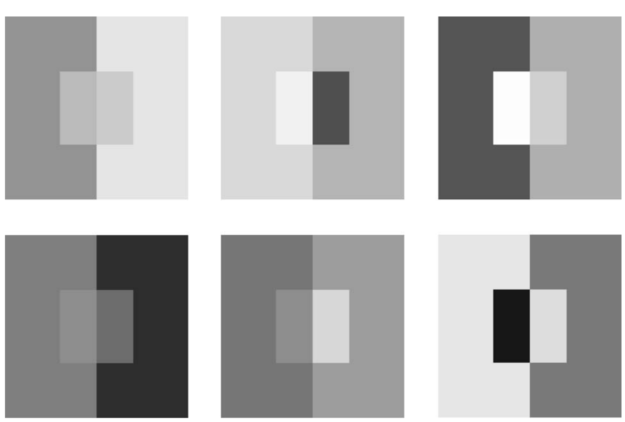

Fig. 5. Examples of cases (from left to right, top to bottom) a translucent Square, overlay with Hole, Right overlay, ambiguous Square or Left overlay, ambiguous Hole or Left overlay, and—finally—a Mosaic, that is to say, a configuration that does not admit of a translucency interpretation. (The omitted cases of ambiguous Square with Right overlay, ambiguous Hole with Right overlay, and the left overlay, are of course simply obtained by holding the page upside down.)

Checking the constraints is a straightforward process. We find that for a random Metelli figure the constraints for a translucent center Square, a translucent center Hole, a translucent Left side, or a translucent Right side have equal probabilities 1/6. (This can be proved through straightforward integration, finding the volume of the parameter space for which the relevant condition is satisfied.) However, the constraints are not necessarily uniquely satisfied. Finding all relevant probabilities is possible through tedious integration over all volumes shown in Fig. 3. However, it is much more convenient to do a simulation and find empirical frequencies. Thus we find that the probability that no translucency condition is possible is 1/2. The ambiguities that may occur are a center square with either a left or a right side, and a center hole with either a left or a right side, the overall probabilities being shown in Table 1.

A notion of how close the various regions in the *XYUV* hypercube are with respect to each other can be obtained if one studies the effect of small perturbations. This is again most easily done through a simulation. In this study we decide with 50% probability on one of the alternatives for the ambiguous cases. For a random perturbation (drawn from normal distributions) of the magnitudes of the luminance levels of 5% we find from a Monte Carlo simulation the probabilities (times 100) shown in Table 2.

Thus the Square has only a very small probability of turning into a Hole and a much higher probability of becoming a Left or Right (simply due to the ambiguous cases, not primarily due to the perturbation) and also a rather high probability of turning into a Mosaic. Similar observations apply to the Hole, Left, and Right categories. The Mosaic category may turn into a Square, Hole, Left, or Right category with equal probability. Thus one expects Square–Hole or Left–Right confusions to be very rare, whereas confusions between any of the translucent categories with the Mosaic (and vice versa) are expected to be fairly frequent.

**Table 1. Probabilities of Various Gestalt Configurations for Random, Uniform Distribution of Image Intensities**

| Configuration | Probability |
|---|---|
| Square | 1/12 |
| Hole | 1/12 |
| Left | 1/12 |
| Right | 1/12 |
| Square and right | 1/24 |
| Square and left | 1/24 |
| Hole and right | 1/24 |
| Hole and left | 1/24 |
| Mosaic | 1/2 |

**Table 2. Frequencies (Times 100) of Various Gestalt Configurations with Perturbation[a]**

| | Square | Hole | Left | Right | Mosaic |
|---|---|---|---|---|---|
| Square | 84 | 0 | 4 | 3 | 6 |
| Hole | 0 | 85 | 5 | 4 | 6 |
| Left | 4 | 5 | 89 | 1 | 3 |
| Right | 3 | 4 | 0 | 88 | 5 |
| Mosaic | 8 | 8 | 9 | 9 | 66 |

[a]Randomly sampled fiducial configurations are perturbed via a 5% normal deviation of the corresponding image intensities (rows were perturbed).

## 3. EXPERIMENT

### A. Methods

In the design of the experiments we had to make a number of more or less arbitrary decisions, which we will discuss here. The basic idea was to have each of a small number of observers categorize a large (though manageable) set of Metelli figures. It is to be expected that a small number of observers will suffice because informal observations suggest that observers are likely to agree in the majority of trials.

A large set of stimuli is required because of the dimensionality of the problem. A Metelli configuration is specified through four numbers in the range of zero to one. Sampling with a (rather coarse) resolution of 10% then yields 10,000 different stimuli. It is evidently out of the question to view them all in a psychophysical session of reasonable duration. We decided to take a random sample out of a uniform distribution. Pilot experiments revealed that an observer can perform a few hundred trials per half hour. Since we also envisaged repeat sessions, we settled on a set of 500 samples.

Pilot experiments revealed that observers consider configurations with very similar intensity levels either trivial or ambiguous. Such instances may be expected to yield little useful information. Thus we decided to consider only configurations in which all four levels were distinct and were multiples of 10% of the maximum intensity. With this coarse resolution all stimuli look indeed like the Metelli configuration, without accidental mergers of areas.

It is probably not realistic to have the observer distinguish cases such as Square and Square or Right. When the stimulus is ambiguous the observer still has to use one of the five categories Square Hole, Left, Right, or Mosaic. Thus if a stimulus in the category Square or Right ends up in either Square or Right the response is counted as veridical.

Random sampling will produce different frequencies of occurrence of the five basic categories. We decided to sample exactly 100 instances of each category for a uniform probability density distribution, subject to the constraints considered above. This should yield a fair and representative sample. The fraction of repeated instances in such a sample is less than 5%.

In order to be able to compare interobserver and intraobserver scatter, we decided to use a single set of 500 instances, revisited 3 times by each observer. The sequence of trials was randomized for any session.

The reflectances are in the range [0,1]. In a scene these reflectances are revealed through the radiance of the beams that are scattered toward the eye. These radiances are proportional to the reflectances, but they also depend on the irradiance of the sheets and the viewing geometry. It is most convenient to refer the radiances to those that would be scattered to the eye by a Lambertian surface of unit albedo in the same spatial attitude (reference white) and by a perfectly black surface (reference black). When presented as an image the Metelli figure is unambiguously perceived only if the reference black and white levels can be visually inferred from the context of the image itself. Notice that reflectances cannot be arbitrarily scaled, as luminances often can. Cutting all luminances by a common factor specifies the same Metelli figure (at least in the photopic domain) only if the white reference is scaled by the same factor. White level anchoring is crucial, as we show in Appendix A.

What this means for experiments in which a monitor is used to present the stimuli is that the Metelli figure has to be presented in a context that visually specifies the white and black references, too. Failure to do this will render the experiment unfit for further analysis because it is incompletely controlled. This is especially a problem with computer presentations, for when the Metelli figures are assembled from scraps of pasted papers the context will usually be present anyway.

In the experiment we presented the Metelli figures on a black-and-white checkered background on a CRT monitor (Fig. 1). The gamma was set to one. The stimulus subtended $12° \times 12°$ of visual angle. Vision was binocular, observers wore their usual correction when applicable. A chin rest was used to fix the head. The room was darkened during the sessions.

### B. Results

We compare the observers responses with the predicted categories. Observers are very consistent in their responses when the stimuli derive from one of the five basic categories. The percentages of correct categorization for the basic categories for the three observers are shown in Table 3.

Thus, with the exception of the Hole category for observer AD (discussed in more detail below), the probabilities of correct categorization are close to 90% or higher.

**Table 3. Percentages of Correct Categorization for Basic Categories for Three Observers**

| Observer | Square | Hole | Left | Right |
|---|---|---|---|---|
| AD | 100 | 2 | 87 | 94 |
| JK | 98 | 90 | 91 | 90 |
| SP | 99 | 98 | 98 | 97 |

In case a variation occurs it is generally a Mosaic response, i.e., the default response. Conversely, the Mosaic category is most often erroneously categorized. This can be seen in the full breakdown per observer (in terms of the absolute number of trials) shown in Tables 4–6.

Notice that only five response categories were available for use by the observer, whereas the stimuli are divided into nine categories. If the mixed cases are divided equally over the corresponding two of the remaining five categories, one obtains (in terms of rounded percentages) the values shown in Tables 7–9. (See also Fig. 6.)

A number of observations can be made *ad oculos* without any need for statistics.

- The responses are very close to the predictions: 64% (AD, but see below), 81% (JK), and 82% (SP) of the responses are veridical.
- The categories Square, Hole, Left, and Right are almost never confused with one another (far below 1%). Only a very small fraction of the responses ends up in a wrong category, with the exception of the Hole category for observer AD.
- Confusions typically arise with the Mosaic category. Since Mosaic is a default category, this is likely to be a failure to categorize, rather than a wrong categorization.
- Stimuli in the Square category were most rarely misclassified.

**Table 4. Responses of Observer AD**

| | Response | | | | |
|---|---|---|---|---|---|
| Stimulus | Square | Hole | Left | Right | Mosaic |
| Square | 258 | 0 | 0 | 0 | 0 |
| Hole | 0 | 4 | 6 | 11 | 234 |
| Left | 33 | 0 | 230 | 0 | 1 |
| Right | 16 | 0 | 0 | 251 | 0 |
| Square or Left | 27 | 0 | 15 | 0 | 0 |
| Square or Right | 21 | 0 | 0 | 6 | 0 |
| Hole or Left | 1 | 0 | 47 | 0 | 0 |
| Hole or Right | 0 | 0 | 0 | 39 | 0 |
| Mosaic | 80 | 1 | 44 | 35 | 140 |

**Table 5. Responses of Observer JK**

| | Response | | | | |
|---|---|---|---|---|---|
| Stimulus | Square | Hole | Left | Right | Mosaic |
| Square | 252 | 0 | 1 | 2 | 3 |
| Hole | 1 | 230 | 0 | 1 | 23 |
| Left | 1 | 0 | 239 | 0 | 24 |
| Right | 0 | 1 | 0 | 241 | 25 |
| Square or Left | 33 | 0 | 8 | 0 | 1 |
| Square or Right | 12 | 0 | 0 | 14 | 1 |
| Hole or Left | 3 | 3 | 36 | 2 | 4 |
| Hole or Right | 0 | 1 | 0 | 32 | 6 |
| Mosaic | 57 | 4 | 31 | 27 | 181 |

**Table 6. Responses of Observer SP**

| | Response | | | | |
|---|---|---|---|---|---|
| Stimulus | Square | Hole | Left | Right | Mosaic |
| Square | 256 | 0 | 0 | 2 | 0 |
| Hole | 1 | 250 | 1 | 2 | 1 |
| Left | 4 | 4 | 258 | 0 | 1 |
| Right | 7 | 0 | 0 | 259 | 1 |
| Square or Left | 24 | 0 | 17 | 1 | 0 |
| Square or Right | 14 | 0 | 0 | 13 | 0 |
| Hole or Left | 1 | 1 | 43 | 3 | 0 |
| Hole or Right | 0 | 2 | 1 | 36 | 0 |
| Mosaic | 40 | 7 | 61 | 60 | 132 |

**Table 7. Collapsed Responses for Observer AD**

| | Square | Hole | Left | Right | Mosaic |
|---|---|---|---|---|---|
| Square | 96 | 0 | 3 | 1 | 0 |
| Hole | 0 | 1 | 10 | 10 | 78 |
| Left | 15 | 0 | 84 | 0 | 0 |
| Right | 9 | 0 | 0 | 91 | 0 |
| Mosaic | 27 | 0 | 15 | 12 | 47 |

**Table 8. Collapsed Responses for Observer JK**

| | Square | Hole | Left | Right | Mosaic |
|---|---|---|---|---|---|
| Square | 94 | 0 | 2 | 3 | 1 |
| Hole | 1 | 78 | 6 | 6 | 9 |
| Left | 6 | 0 | 84 | 0 | 9 |
| Right | 2 | 0 | 0 | 88 | 10 |
| Mosaic | 19 | 1 | 10 | 9 | 60 |

**Table 9. Collapsed Responses for Observer SP**

| | Square | Hole | Left | Right | Mosaic |
|---|---|---|---|---|---|
| Square | 94 | 0 | 3 | 3 | 0 |
| Hole | 1 | 84 | 8 | 7 | 0 |
| Left | 5 | 1 | 92 | 1 | 0 |
| Right | 5 | 0 | 0 | 94 | 0 |
| Mosaic | 13 | 2 | 20 | 20 | 44 |

- Stimuli in the Hole category were more often misclassified, observer AD being an extreme case with 98%, though qualitatively similar to observers JK (22)% and SP (15%).
- Misclassified stimuli in the Mosaic category were more often put in the Square category than in any other (27%, 19%, and 13%).
- Misclassified stimuli in the Mosaic category were least likely to be put in the Hole category (0%, 1%, and 2%).
- In the mixed categories involving the Square, the response is much more likely to be Square than Left or Right (70%, 65%, and 55%).
- In the mixed categories involving the Hole, the response is much less likely to be Hole than Left or Right (1%, 8%, and 5%).

Thus observer JK and (to a lesser degree) SP evidently prefer the Square over the Hole category, an effect that occurs in extreme form with observer AD. It is clearly due to the fact that smaller, convex regions are more easily seen as figure than are global, nonconvex areas. Otherwise the pattern of confusions is remarkably like that we find for the slightly perturbed model.

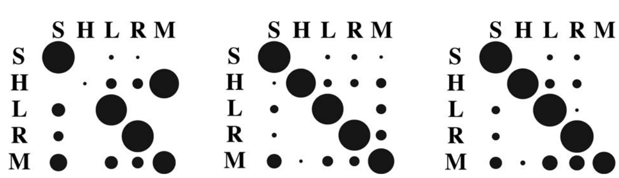


Fig. 6. Collapsed counts of observers (left to right) AD, JK, and SP.

In a Monte Carlo simulation we set a noise level that generated the level of confusion empirically found for the Mosaic category for the average observer in our experiment. The required noise level amounted to 5% of the maximum intensity. Then we counted for each stimulus how many times a confusion occurred. This list of a hundred numbers was correlated with a similar list derived from the data. The Pearson correlation coefficient is 0.81 (AD), 0.82 (JK), 0.76 (SP), whereas it is about 0.8 for independent, repeated simulations. The observers thus commit similar errors; the interobserver correlations were 0.85 (AD–JK), 0.88 (AD–SP), and 0.80 (JK–SP). The simulation effectively selects sensitive cases where a small perturbation would throw the category into a different bin. As it turns out, the observers frequently confuse exactly these sensitive cases.

This can be seen especially clearly in Fig. 7. The simulation was repeated 100 times for each of the 100 stimuli, yielding a good estimate of the probability of shifting categories. The instances were sorted with respect to their susceptibilty to perturbation. For the observers the empirical probabilities are 0, 1/3, 2/3, or 1. Evidently, the observers frequently commit errors for the sensitive instances but almost never for the robust ones. The model accounts very well for the probabilities of miscategorizations committed by the observers.

The model can easily account for the data if we assume a 5% noise level and moreover assume that

- Hole classifications are not accepted (are put in the default Mosaic category) with a probability that is 0% for SP, 10% for JK, and 100% for AD;

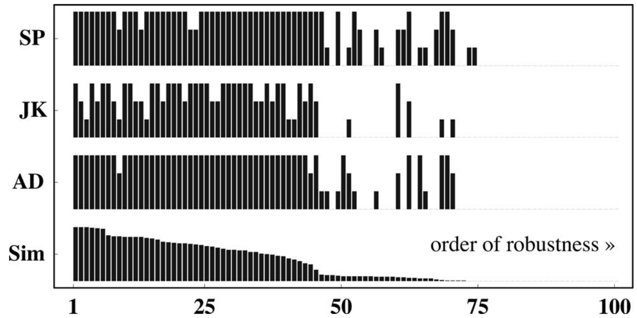


Fig. 7. In a simulation the actual stimuli used in the experiment were perturbed 100 times and the perturbed instance categorized. This yields a probability of miscategorization plotted as the "Sim" bars. The stimuli have been sorted by this probability. The empirical probability of miscategorization by the three observers AD, JK, and SP has also been plotted in the same sequence (here the resolution is only 1/3 because there were only three repeated trials per stimulus). We show the simulated and empirical frequencies for the miscategorization of Mosaic stimuli. Notice that sensitive stimuli in the simulation are frequently miscategorized by the observers, and robust instances hardly ever. The stimuli that were never miscategorized in the simulation were never confused by the observers, either.

- Ambiguous Square or Left–Right categorizations are put in the Square category with 75% probability;
- Ambiguous Hole or Left–Right categorizations are put in the Left or Right categories (whichever applies).

Apparently the model describes the responses remarkably well; the confusions committed by an observer can be largely accounted for by a small noise level and a bias in favor of Squares and against Holes. The exceptions are in accord with well-known Gestalt principles.

## 4. DISCUSSION

We started this work because of our informal observation that observers spontaneously report perceptions that do not fit the conventional framework. Instead of curing this in the usual way, that is by setting up a forced choice task with an *a priori* prescribed dichotomous choice, we attempted a more general analysis and offered our observers a wider choice of categories [35,36].

Such a method has been pioneered by Kitaoka [10], although the categories considered in that study appear to have been selected on an *ad hoc* basis. In this study the observers effectively ignored some of the categories offered to them, which is a clear indication that the choice of categories was less than satisfactory. Nevertheless, this is an exemplary and original study, going beyond the conventional paradigm.

An important problem that arises in these experiments is the high dimensionality of the parameter space. There exist (at least) four independent parameters; the parameter space is a four-dimensional (unit) hypercube. In order to sample the parameter space uniformly and densely, one needs a very large number of instances. For instance, in the Kitaoka [10] study mentioned earlier only 50 instances were used, effectively only 2 or 3 samples per dimension, an evident case of severe undersampling. The regions in parameter space belonging to single perceptual categories are small and numerous, requiring many hundreds of samples to resolve.

With our (admittedly overly simplified) analysis we arrive at five generic categories. The simplification (the number of possible interpretations of the Metelli configuration is very much higher than that) is arrived at through the application of basic Gestalt principles. In the experiment our observers had no difficulties in categorizing the instances, and the category counts closely reflect the theoretical probabilities.

Since the categories used in the experiment are different from those generally used in similar studies, it is not easily possible to attempt in-depth comparisons with published results. We believe it likely that our Mosaic category is similar to the conventional no translucency case, our Square category to the conventional translucency case, and our Left and Right categories to the (not often used) partial translucency case, whereas our Hole category is hard to place. Perhaps the hole category is analogous to the no translucency, perhaps to the partial translucency case. The existence of real (i.e., theoretically predicted) ambiguous cases is—to the best of our knowledge—not recognized in the literature. Such cases are often recognized by observers and may then be denoted multiple translucency.

We find (detailed in Appendix A) that the predictions of the Metelli constraints are very similar to those of the (more ecologically valid) Kubelka–Munk analysis, the differences being sufficiently minor that one might not be able to detect them in psychophysical experiments. The Metelli constraints are evidently sufficient to account for our results. Even the (infrequent) confusions are explained in satisfactory, quantitative detail.

The fact that even small perturbations of the lightnesses of the four areas in the Metelli configuration may suffice to change the Gestalt category is not recognized in the literature. We consider this problematic, especially in view of the fact that the total number of instances used in typical studies is very small, given the high dimensionality of parameter space.

The asymmetry between the responses in the Square and the Hole categories found in the psychophysical results are not predicted by the theory. One clearly has to add additional assumptions. That smallish, coherent objects are more frequent than smallish holes in larger coherent objects is, of course, ecologically evident. The asymmetry may have to do with fundamental rules of figure–ground assignment. Idiosyncratic differences might be speculated to correlate with other visual functions that subserve the perception of depth order in natural circumstances, such as the nature of an individual's binocularity (presence or absence of certain disparity pools). This is an interesting issue for further research.

If the actual research interest is in the perception of translucency, rather than Gestalt interpretation combined with perceptual translucency, it makes much sense to substitute the simple X-crossing configuration (see Fig. 8) for the conventional Metelli configuration. For this simple configuration there is a Mosaic interpretation (no perceptual translucency) and interpretations with Left, Right, Lower, or Upper translucent overlays. These five

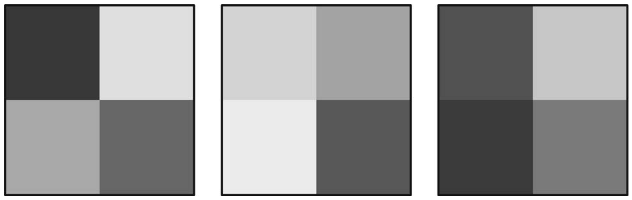

Fig. 8. Left to right, Mosaic, an Upper, and ambiguous Lower and Left case for the X-crossing configuration.

response categories do not suffer from the Hole–Square asymmetry. The theoretical analysis of this case is in all respects similar to that of the Metelli configuration. For this configuration the partial translucency category is not present, but cases of multiple translucency occur. The ambiguous cases appear as combinations of, e.g., Left and Lower, and so forth, translucent overlays. In a uniform sampling the frequency of finding a Mosaic is 1/2 and the four pure overlay cases (Left, etc.) occur with frequencies 1/12, whereas the four ambiguous cases (Lower and Left, etc.) occur with frequencies 1/24. In an experiment one would probably prefer a stratified sampling method like the one used in this study.

Where comparable, our results agree well with the existing literature on the Metelli configuration *per se*, but of course this work applies only to the classic Metelli paradigm. The modern literature has progressed far beyond this [37–39], but the present work can hardly be extrapolated beyond the conventional paradigm. We believe discrepancies with earlier work to be due mainly to these factors:

- Lack of explicit anchoring of stimulus levels,
- Overly restrictive response categories,
- Undersampling of the parameter space,
- Failure to take stimulus noise immunity into account.

All of these factors are crucial (detailed in Appendix A), yet have never been considered in a single study. In most of the early literature on the Metelli case the anchoring of the white level is ill defined, and in many cases so are the response categories. This makes it very hard indeed to compare our results with those of earlier authors.

In summary, we conclude that Metelli's optical constraints have excellent predictive power when properly combined with an analysis of the possible Gestalt categories provided by the geometrical configuration.

## APPENDIX A

### 1. Preliminaries

We consider the standard psychophysical setup where one has two background colors $\mathcal{X}$ and $\mathcal{Y}$ and an overlay composed of two colors $\mathcal{P}$ and $\mathcal{Q}$. In the case that a physical interpretation is feasible the overlay will be characterized by the same physical parameters (e.g., turbidity, layer thickness, …) for both colors $\mathcal{P}$ and $\mathcal{Q}$. Likewise, in a cognitive model the colors $\mathcal{P}$ and $\mathcal{Q}$ are assumed to be transformations of the colors $\mathcal{X}$ and $\mathcal{Y}$ by a common cause, that is, the same transformation. The transformation in a cognitive model is not necessarily explained through physical causes (a physical interpretation need not exist), although this often will be the case.

For a start we consider only gray values; i.e., the colors are fully characterized by their gray values and will be specified by a number in the range [0,1], 0 standing for $\mathcal{K}$, black (K refers to "key") and 1 for $\mathcal{W}$, white. In a physical model the gray values are also interpreted as reflectances.

In order to study the problem formally we map the pairs in the bicolor diagram, the unit square, in which the axes parameterize the gray values of the left- and right-hand colors. Thus both the pair $\mathcal{X}$ and $\mathcal{Y}$ and the pair $\mathcal{P}$ and $\mathcal{Q}$ are represented as points within the square (see Fig. 9).

Consider a background $\{\mathcal{X},\mathcal{Y}\}$. What are possible locations for the overlay $\{\mathcal{P},\mathcal{Q}\}$? Clearly all opaque, uniform overlays are possible, so the diagonal $\{\mathcal{K},\mathcal{K}\}$–$\{\mathcal{W},\mathcal{W}\}$ is part of the region. So is the point $\{\mathcal{X},\mathcal{Y}\}$, because it corresponds to an empty overlay. It is intuitively evident that gray values cannot cross over; thus the region is bounded by the diagonal. Moreover, it is intuitively obvious that a translucent overlay cannot amplify the contrast; thus the lines $\{\mathcal{K},\mathcal{K}\}$–$\{\mathcal{X},\mathcal{Y}\}$ and $\{\mathcal{W},\mathcal{W}\}$–$\{\mathcal{X},\mathcal{Y}\}$ are also part of the boundary. The locus of possible overlays will thus be part of the convex hull of the diagonal and the background point. Both the diagonal and the background point will be on the boundary of the region.

### 2. Metelli Model

Metelli introduced a simple model that could be implemented in the laboratory by using the episcopister but that can be interpreted more generally in terms of some luminous atmosphere. We regard it as a phenomenological model. Metelli introduces the expression

$$\mathcal{P} = \alpha\mathcal{X} + (1-\alpha)\mathcal{T}, \tag{A1}$$

$$\mathcal{Q} = \alpha\mathcal{Y} + (1-\alpha)\mathcal{T}, \tag{A2}$$

where $\mathcal{T}$ is a color characteristic of the luminous atmosphere and $\alpha$ a parameter that varies in the range [0,1] that summarizes the total turbidity.

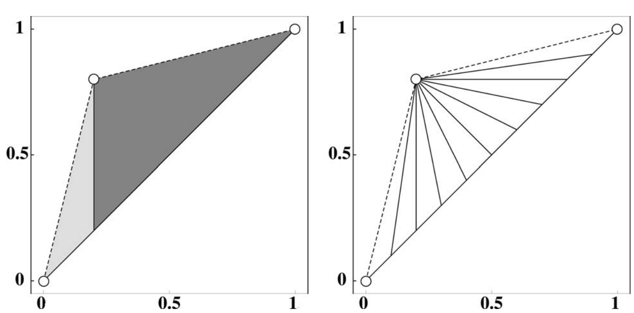


Fig. 9. Left, unit square of color pairs $\{\mathcal{X},\mathcal{Y}\}$ and $\{\mathcal{P},\mathcal{Q}\}$. The indicated points are the black point $\{\mathcal{K},\mathcal{K}\}=\{0,0\}$, the white point $\{\mathcal{W},\mathcal{W}\}=\{1,1\}$, and the back ground $\{\mathcal{X},\mathcal{Y}\}$. The shaded region indicates the region where the overlay pair has a valid Metelli interpretation; the shades indicate partial constraints. Right, the opaque overlay limit lies on a diagonal point $\{\mathcal{T},\mathcal{T}\}$. We picked many such diagonal points and interpolated with the background via the parameter $\alpha$, thus generating the (straight) orbits. Notice that the orbits sweep out the convex hull of the diagonal and the background.

Notice that the model is "conservative" (conserves radiant power). A Metelli-type model that includes absorption would be

$$\mathcal{P} = \alpha\mathcal{X} + \beta\mathcal{T}, \tag{A3}$$

$$\mathcal{Q} = \alpha\mathcal{Y} + \beta\mathcal{T}, \tag{A4}$$

with $\alpha+\beta \leqslant 1$. For $\alpha+\beta=1$ you regain the Metelli model; otherwise there is an additional absorption. We will analyze the original Metelli model here.

Given the two pairs of colors $\{\mathcal{X},\mathcal{Y}\}$ and $\{\mathcal{P},\mathcal{Q}\}$, one can solve for the color of the overlay $\mathcal{T}$ and the Metelli parameter $\alpha$:

$$\mathcal{T} = \frac{\mathcal{X}\mathcal{Q} - \mathcal{Y}\mathcal{P}}{(\mathcal{X}-\mathcal{Y}) - (\mathcal{P}-\mathcal{Q})}, \tag{A5}$$

$$\alpha = \frac{\mathcal{P}-\mathcal{Q}}{\mathcal{X}-\mathcal{Y}}. \tag{A6}$$

There exists a Metelli interpretation of the configuration if $0 \leqslant \mathcal{T} \leqslant 1$ and $0 \leqslant \alpha \leqslant 1$. Consider (without loss of generality) that $\mathcal{X} \leqslant \mathcal{Y}$. Then the constraints can be reduced to the following conditions.

- $\mathcal{P} \leqslant \mathcal{Q}$, that is to say, the contrast never reverses polarity.
- If $\mathcal{P}$ is darker than $\mathcal{X}$, then $\mathcal{P}/\mathcal{Q} \leqslant \mathcal{X}/\mathcal{Y}$; that is to say, the contrast of the overlay can never exceed that of the ground. For $\{\mathcal{P},\mathcal{Q}\}=\mu\{\mathcal{X},\mathcal{Y}\}$ one obtains the equality; thus the constraint is bounded by the line $\{\mathcal{X},\mathcal{Y}\}-\{\mathcal{K},\mathcal{K}\}$.
- If $\mathcal{P}$ is lighter than $\mathcal{X}$, then $\mathcal{Y}\mathcal{P}-\mathcal{X}\mathcal{Q} \leqslant (\mathcal{P}-\mathcal{Q})-(\mathcal{X}-\mathcal{Y})$, which is a linear constraint. For $\{\mathcal{P},\mathcal{Q}\}=\mu\{\mathcal{X},\mathcal{Y}\}+(1-\mu)\{\mathcal{W},\mathcal{W}\}$ one obtains the equality; thus the constraint is bounded by the line $\{\mathcal{X},\mathcal{Y}\}-\{\mathcal{W},\mathcal{W}\}$.

Consequently, these constraints simply coincide with the convex region outlined above (Fig. 9, left).

When you let the Metelli parameter run from zero to one, the pair $\{\mathcal{P},\mathcal{Q}\}$ moves via a rectilinear path from $\{\mathcal{T},\mathcal{T}\}$ on the diagonal of the bicolor diagram to the point $\{\mathcal{X},\mathcal{Y}\}$ (Fig. 9, right).

### 3. Importance of Luminance Anchoring

Suppose that one would not use white–black anchoring as in Fig. 1; how would the possible interpretations change? Then any luminance level could be assigned to white, whereas the black level would remain unchanged (we ignore foggy circumstances here). Given the freedom to assign the white level, we obtain a one-parameter family of possible interpretations. These possibilities might in principle serve to increase the ambiguity.

Numerical investigation of 10,000 random cases shows that there exist Mosaic configurations that turn into either a Square or Right-bar, Square or Left-bar, Hole or Right-bar, or Hole or Left-bar configuration when the white level is set higher than the fiducial white level, though not all Mosaic cases are of this kind. Symbolic manipulation of the logical expressions (using a symbolic algebra package, the expressions are rather complicated) serves to verify this conclusion formally.

Consider an example: assume $\mathcal{X} > \mathcal{Y}$ and a fiducial white level of unity. Then one type of the Mosaic category (there exist various other possibilities) is characterized through

$$(\mathcal{P} \geqslant \mathcal{Q}) \wedge (\mathcal{Q}\mathcal{X} \geqslant \mathcal{P}\mathcal{Y}) \wedge (\mathcal{P} + \mathcal{Q}\mathcal{X} + \mathcal{Y} \geqslant \mathcal{Q} + \mathcal{X} + \mathcal{P}\mathcal{Y}) \wedge (\mathcal{X} < \mathcal{P}), \tag{A7}$$

whereas the Square or Left category is characterized through

$$(\mathcal{P} > \mathcal{Q}) \wedge (\mathcal{X} < \mathcal{P}) \wedge (\mathcal{Y} < \mathcal{Q}) \wedge (\mathcal{P} + \mathcal{Q}\mathcal{X} + \mathcal{Y} < \mathcal{Q} + \mathcal{X} + \mathcal{P}\mathcal{Y}). \tag{A8}$$

Suppose the first constraint is TRUE, then the second evaluates to FALSE. Now assume a different white level $\xi > 1$ (say). Substituting

$$\mathcal{X} \to \xi^{-1}\mathcal{X}, \quad \mathcal{Y} \to \xi^{-1}\mathcal{Y}, \quad \mathcal{P} \to \xi^{-1}\mathcal{P}, \quad \mathcal{Q} \to \xi^{-1}\mathcal{Q}, \tag{A9}$$

we find that the second constraint evaluates to TRUE if

$$\xi > \frac{\mathcal{Q}\mathcal{X} - \mathcal{P}\mathcal{Y}}{(\mathcal{X}-\mathcal{Y}) - (\mathcal{Q}-\mathcal{P})} > 1. \tag{A10}$$

Thus the nominal Mosaic configuration can be interpreted as a Square or Left configuration (thus either Square or Left would be correct responses) simply by assuming a higher white level. Thus the anchoring of the white level is of crucial importance.

This has important consequences for psychophysics, since it shows that the anchoring is absolutely necessary. It renders much of the literature data difficult to interpret. In classical studies, where a forced choice between transparent (meaning a translucent square in front of a bipartite background) and not transparent (anything else) is required, there will be ambiguous stimuli if the observer is free to assume a white level *ad libitum*.

### 4. Kubelka–Munk Approximation

A convenient approximation for the optical properties of turbid layers is provided by the two-flux model introduced by Schuster in the early twentieth century for the optics of planetary atmospheres and nowadays generally denoted the Kubelka–Munk approximation.

Consider a translucent layer in the two-flux approximation. We ignore Fresnel reflection or refraction at the interfaces; i.e., a piece of translucent paper is a good implementation. Moreover, we assume that absorption can be neglected, that is to say, we assume a conservative system. The latter assumption is by no means necessary, but it closely mimics the (conservative) Metelli model. The two-flux model ignores lateral diffusion and thus cannot account for blurring of backgrounds seen through a turbid layer. It is easy enough to incorporate such effects, but the two-flux model perfectly captures the Metelli case.

The relevant Kubelka–Munk solution is

$$\mathcal{P} = \frac{\dfrac{\mathcal{X}-\mathcal{T}}{\mathcal{T}} - \mathcal{T}\left(\mathcal{X} - \dfrac{1}{\mathcal{T}}\right)\exp\left[\alpha\left(\dfrac{1}{\mathcal{T}} - \mathcal{T}\right)\right]}{\mathcal{X} - \mathcal{T} - \left(\mathcal{X} - \dfrac{1}{\mathcal{T}}\right)\exp\left[\alpha\left(\dfrac{1}{\mathcal{T}} - \mathcal{T}\right)\right]}, \tag{A11}$$

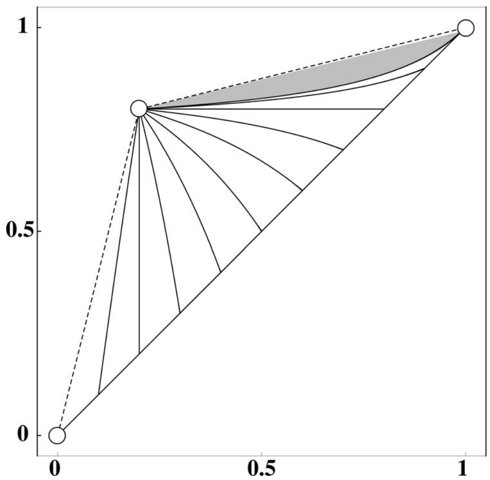


Fig. 10. The orbits in the bicolor diagram for the Kubelka–Munk approximation are curved. They do not exhaust the convex region covered in the Metelli model; in the shaded region one has a valid Metelli interpretation but no Kubelka–Munk interpretation.

$$\mathcal{Q} = \frac{\dfrac{\mathcal{Y}-\mathcal{T}}{\mathcal{T}} - \mathcal{T}\left(\mathcal{Y}-\dfrac{1}{\mathcal{T}}\right)\exp\left[\alpha\left(\dfrac{1}{\mathcal{T}}-\mathcal{T}\right)\right]}{\mathcal{Y}-\mathcal{T}-\left(\mathcal{Y}-\dfrac{1}{\mathcal{T}}\right)\exp\left[\alpha\left(\dfrac{1}{\mathcal{T}}-\mathcal{T}\right)\right]}, \qquad \text{(A12)}$$

where the parameters have a meaning similar to that in the Metelli model: the color $\mathcal{T}$ is the color of an infinitely thick (opaque) layer, and the coefficient $\alpha$ represents the turbidity, that is, the product of the scattering coefficient of the medium and the thickness of the layer. The color $\mathcal{T}$ is a material constant that depends on the absorption and scattering coefficients of the medium. For our purposes the turbidity $\alpha$ is simply a parameter in the range $(0,\infty)$.

The Kubelka–Munk model is very similar to the Metelli model. The essential difference is that the Kubelka–Munk model yields curved orbits instead of straight ones (see Fig. 10) because the Kubelka–Munk model is nonlinear.

As a result of the nonlinearity the permitted region for the Kubelka–Munk model lies indeed within the convex hull $[\{\mathcal{K},\mathcal{K}\},\{\mathcal{X},\mathcal{Y}\},\{\mathcal{W},\mathcal{W}\}]$, but does not exhaust it. Thus the Kubelka–Munk model is somewhat more restrictive than the Metelli model. The difference (shaded region in Fig. 10) is not large, though, and may not be detectable with psychophysical experiments—assuming that the human observer only perceives physically possible structures as translucent, of course. (Fig. 10.)

It is not possible to give simple analytical expressions for the Kubelka–Munk limits, but it is easy enough to check whether a certain configuration has a valid Kubelka–Munk interpretation numerically. One simply computes the Kubelka–Munk parameters $\mathcal{T}$ and $\alpha$ from the color pairs $\{\mathcal{X},\mathcal{Y}\}$ and $\{\mathcal{P},\mathcal{Q}\}$ and checks whether $0\leqslant\mathcal{T}\leqslant 1$ and $\alpha\geqslant 0$.

## ACKNOWLEDGMENTS

This work was sponsored via the European program Visiontrain, contract MRTNCT2004005439. Sylvia Pont was supported by the Netherlands Organisation for Scientific Research (NWO). We thank Robert Volcic for helping us find obscure references.

## REFERENCES AND NOTES

1. W. Metzger, *Gesetze des Sehens* (Waldemar Kramer, 1975).
2. *Laws of Seeing [Gesetze des Sehens]*, L. Spillmann, M. Wertheimer, and S. Lehar, translators (MIT Press, 2006).
3. F. Metelli, "Zur Analyse der phänomenalen Durchsichtigkeitserscheinungen," in *Gestalt und Wirklichkeit*, R. Müller and J. Fischl, eds. (Duncker & Humblot, 1967), pp. 258–304.
4. F. Metelli, "An algebraic development of the theory of perceptual transparency," Ergonomics **13**, 59–66 (1970).
5. F. Metelli, "Achromatic colour conditions in the perception of transparency," in *Perception: Essays in Honor of James J. Gibson*, R. B. MacLeod and H. L. Pick, eds. (Cornell U. Press, 1974), pp. 95–116.
6. F. Metelli, "The perception of transparency," Sci. Am. **230**(4), 91–98 (1974).
7. In the literature, called "partial transparency." The concept was introduced by Beck in 1984 [J. Beck, K. Prazdny, and R. Ivry, "The perception of transparency with achromatic colors," Percept. Psychophys. 35, 407–422 (1984)]. See also [8,9]. The concept of "partial transparency" is not very well defined; we believe it to correspond to our Left or Right overlay categories.
8. J. Beck and R. Ivry, "On the role of figural organization in perceptual transparency," Percept. Psychophys. **44**, 585–594 (1988).
9. F. Metelli, O. da Pos, and A. Cavedon, "Balanced and unbalanced, complete and partial transparency," Percept. Psychophys. **38**, 354–366 (1985).
10. A. Kitaoka, "A new explanation of perceptual transparency connecting the X-junction contrast-polarity model with the luminance-based arithmetic model," Jpn. Psychol. Res. **47**, 175–187 (2005).
11. W. Gerbino, C. I. F. H. J. Stultiens, J. M. Troost, and Ch. M. M. de Weert, "Transparent layer constancy," J. Exp. Psychol. **16**, 3–20 (1990).
12. S. C. Masin, "The luminance conditions of transparency," Perception **26**, 39–50 (1997).
13. M. Singh and D. D. Hoffman, "Part boundaries alter the perception of transparency," Psychol. Sci. **9**, 370–378 (1998).
14. M. Singh and B. L. Anderson, "Towards a perceptual theory of transparency," Psychol. Rev. **109**, 492–519 (2002).
15. R. W. Fleming and B. L. Anderson, "The perceptual organization of depth," in *The Visual Neurosciences*, L. Chalupa and J. S. Werner, eds. (MIT Press, 2003), pp. 1284–1299.
16. P. Kubelka and F. Munk, "Ein Beitrag Zur Optik der Farbanstriche," Z. Tech. Phys. (Leipzig) **12**, 593–601 (1931).
17. K. Koffka, *Principles of Gestalt Psychology* (Lund Humphries, 1935).
18. M. Wertheimer, "Über Gestalttheorie," an address before the Kant Society, Berlin, 7th December, 1924, Erlangen, 1925, translation by W. D. Ellis in *Source Book of Gestalt Psychology* (Harcourt, Brace, 1938).
19. R. Arnheim, *Visual Thinking* (U. California Press, 1969).
20. W. Kohler, *Gestalt Psychology* (Liveright, 1947).
21. G. Kanizsa, *Organization in Vision: Essays on Gestalt Perception* (Praeger, 1979).
22. F. Metelli, "Oggettualità stratificazione e risalto nell'organizzazione (percettiva) di figura e sfondo," Arch. Psicol Neurol. Psichiatr **II**(4), 831–841 (1941).
23. F. Metelli, "Morfologia dei fenomeni di completamento nella percezione visiva," in *Gestalthaftes Sehen*, E. von Ehrenfels and F. Weinhandl, eds. (Wissenschaftliche Buchgesellschaft, 1960), pp. 266–278.
24. F. Metelli, "On the visual perception of transparency," in

*Studies in Perception: Festschrift for Fabio Metelli*, G. B. F. D'Arcais, ed. (Martello Giunti, 1975), pp. 488–496.

25. F. Metelli, "Shadows without penumbra," in *Gestalttheorie in der modernen Psychologie (Metzger Festschrift)*, S. Ertal, L. Kemmier, and S. Michael, eds. (Steinkopff, 1975), pp. 200–209.
26. F. Metelli, "What does 'more transparent' mean? A paradox," in *Vision and Artifact*, M. Henle, ed. (Springer, 1976), pp. 19–24.
27. F. Metelli, O. da Pos, and A. Cavedon, "Some conditions regarding localization and mode of appearance of achromatic colors," Atti Mem. Accad. Pata. Sci. Lett. Arti **89**, 213–221 (1976–1977).
28. F. Metelli, S. C. Masin and A. M. Manganelli, "Partial transparency," Atti Accad. Patav. Sci. Lett. Arti 1979–1980, Cl. Sci. morali, Lett. Arti XVII (Parte III), 115–169 (1981).
29. F. Metelli, "Stimulation and perception of transparency," Psychol. Res. **47**, 185–202 (1985).
30. F. Metelli, "Osservazioni sull'articolo di Marco Sambin: trasparenza con superfici indotte'," G. Ital. Psicolo. **XI**, 393–398 (1984).
31. F. Metelli, "Su alcune condizioni spazio-figurali della trasparenza," in *Conoscenza e Struttura*, W. Gerbino, ed. (Il Mulino, 1985), pp. 303–331.
32. D. Hoffman, *Visual Intelligence: How We Create What We See* (Norton, 1998).
33. M. Albert, D. Hoffman, "Genericity in spatial vision," in *Geometric Representations of Perceptual Phenomena: Articles in Honor of Tarow Indows on His 70th Birthday*, R. D. Luce, M. D. D'Zmura, D. Hoffman, G. J. Iverson, and A. K. Romney, eds. (Erlbaum, 1995), pp. 95–112.
34. D. B. Gauld, *Differential Topology: An Introduction* (Marcel Dekker, 1982).
35. Metelli, who was an acute visual observer, remarks on our Hole category in a footnote: "In questo caso [namely $|u-v| > |x-y|$] si determina un effetto curioso …. Invece delle regioni centrali diventano trasparenti, e cioè assumono le funzioni di U e V, le regioni laterali, e quindi si restaura la condizione necessaria [$|u-v| < |x-y|$" [expressions in our notation] but does not appear to have used this insight in the way we did. See footnote 21, p. 171 in F. Metelli, "Stimolazione e percezione della trasparenza con tonalità acromatiche," Ric. Psicol. **26**, 159–180 (1983). Percepts related to holes have been more widely studied in binocular vision, e.g., see [36].
36. B. L. Anderson and B. Julesz, "A theoretical analysis of illusory contour formation in stereopsis," Psychol. Rev. **102**, 705–743 (1995).
37. B. L. Anderson, M. Singh, and J. Meng, "The perceived transmittance of inhomogeneous surfaces and media," Vision Res. **46**, 1982–1995 (2006).
38. R. Robilotto and Q. Zaidi, "Perceived transparency of neutral density filters across dissimilar backgrounds," J. Vision **4**(3), 183–195 (2004).
39. R. Kasrai and F. A. A. Kingdom, "Precision, accuracy, and range of perceived achromatic transparency," J. Opt. Soc. Am. A **18**, 1–11 (2001).