

## Compositorial ‘Weight’ & ‘Luminance’

Jan Koenderink<sup>1,2,\*</sup>, Andrea van Doorn<sup>2</sup> and Karl Gegenfurtner<sup>3</sup>

<sup>1</sup>Experimental Psychology, University of Leuven (KU Leuven), Belgium

<sup>2</sup>Experimental Psychology, Utrecht University, The Netherlands

<sup>3</sup>Abteilung Allgemeine Psychologie, Justus-Liebig Universität, Giessen, Germany

Received 19 September 2016; accepted 14 June 2017

---

### Abstract

Compositorial weight might be understood as an operational definition of salience. It is not a psychophysical entity, but holds a key position between psychophysics and aesthetics. Several factors ranging over raw photometric/colorimetric parameters, various kinds of psychophysical contrast, image geometry, even semantic properties are readily shown to influence weight. A down-to-earth proposition is that luminance might play a dominant role. We investigate this notion and show that luminance *per se* is hardly important, except in certain paradigms like the ones considered here. We find that observers indeed readily judge weight based on luminance in such paradigms, although there are strong idiosyncratic differences. Our results have some generic implications for graphical design.

### Keywords

Art, phenomenology, perception, composition, luminance

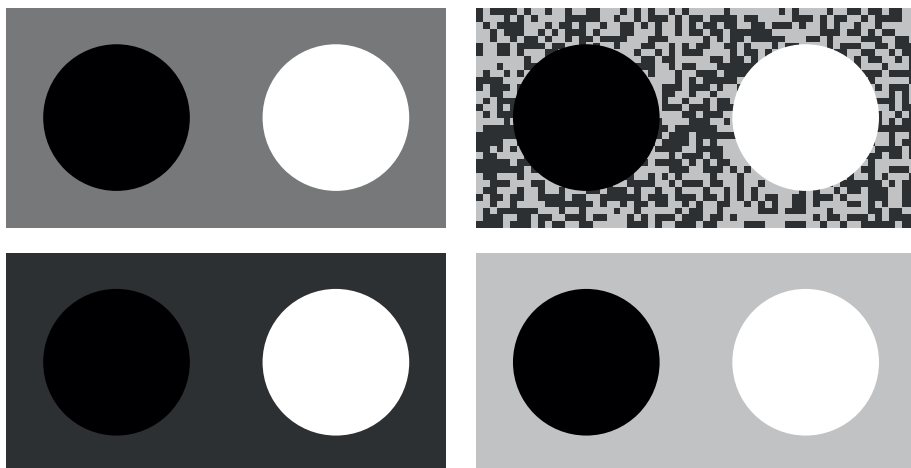
## 1. Introduction

In experiments involving Gestalt factors it often appears that the photometric/colorimetric notion of ‘luminance’ is not an effective predictor for the intuitive ‘balancing’ of differently colored patches in a composition. The extreme case illustrating this involves the balancing of the ‘weight’ of a white patch and a black patch on a common mid-tone gray background (Fig. 1, top left). Such balancing is—among other things—affected by size, edge-quality, shape and position in the picture (Dondis, 1973; Graves, 1951; Taylor, 1964; Wong, 1993). Here luminance *per se* is evidently irrelevant.

Examples from art education do not necessarily involve color or luminance (the latter roughly corresponding to the ‘tone’ or ‘value’ as used by the artist)

---

\*To whom correspondence should be addressed. E-mail: koenderinkjan@gmail.com

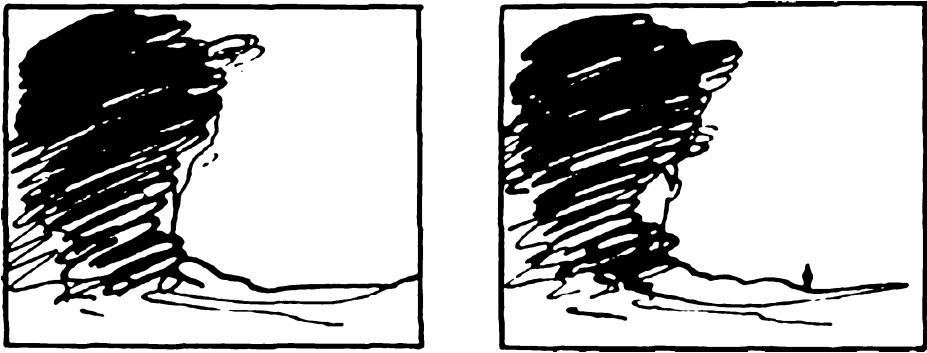


**Figure 1.** Four instances of simple compositions with two blobs on (perhaps statistically) uniform grounds. The two patches are identical in all cases. Are these ‘balanced’ configurations, or do they look lopsided, verging towards left or right? One’s intuitive decision will be influenced by many factors, such as the shades of gray, the contrasts with the background, the sizes, the position in the frame, the edge quality and so forth. Of course, there are also idiosyncratic factors in play. In the top-right example the background has no particular shade of gray, in the top left it has a uniform tone yielding a balanced composition. In the examples at bottom the design is unbalanced due to a different choice of background tone. At left, it veers towards the white blob, at right towards the black blob.

at all. Figure 2, from a treatise on artistic technique, shows an example that has little relation to either tone or color. Thus compositorial ‘weight’ may derive from *very* different qualities than the mere photometric or colorimetric ones. We have shown this before in the chromatic domain (Koenderink *et al.*, in press). This paper also adds additional technical (colorimetric) specifications.

We consider ‘weight’ in the compositorial sense (Locher, Overbeeke and Stappers, 2005; MacManus, Edmondson and Rodger, 1985; Mokaran, 2007; Monroe, 1926; Morriss, Dunlap and Hammond, 1982; Parada-Castellano, 2016; Pinkerton and Humphrey, 1974; Wise and Wise, 1988; Wright, 1962). It is perhaps like an operational salience measure. Observers simply use their eye measure, rather than photometric or colorimetric concepts.

From the perspective of colorimetry one may be inclined to propose *luminance* as a (at least potentially) relevant weight factor. On a generic display unit the red, green and blue channels have luminances in (roughly) the ratios  $R:G:B = 3:6:1$  (CIE, 1990; Eisner and MacLeod, 1980; Smith and Pokorny, 1987). Since luminance is a linear functional of spectral radiant power density, this implies that a bright blue B will be equiluminant with a ‘yellow’  $0.1(R+G)$  (typically  $W=R+G+B$  looks white). This will never be



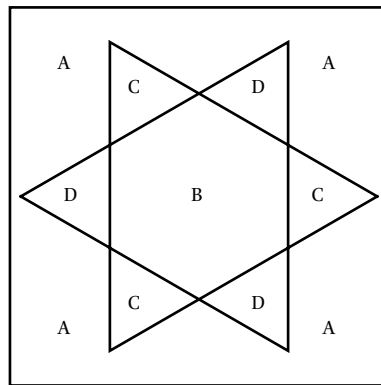
**Figure 2.** An example (Diagram XXVII: ‘Illustration How Interest May Balance Mass’) by Harold Speed (1913). Notice how unbalanced the landscape composition at left appears. In the sketch at right the introduction of a tiny human figure manages to effectively ‘balance’ the group of trees compositionally.

accepted as at all reasonable by artists since the ‘dark yellow’ will appear (at best) like a dark drab brown, not ‘yellow’ at all (Buck, 2014), so it will *never* balance against the bright blue.

In a recent study (Koenderink *et al.*, in press) we showed that the closest ‘balance’ between the primary colors for graphical applications appears to be more like  $R:G:B=1:1:1$ , the so-called ‘equipollent’ condition. Indeed, the cardinal colors appear to mutually balance each other, implying that the ‘weight’ functional is like the maximum of their R, G and B parts. In practice, all display units conform to this, no doubt because display technology is driven by customer demand. It echoes Schopenhauer’s (1870) theory of colors as proper ‘parts of white’, rather than conforming to CIE luminance. Indeed, a monitor with equiluminant color channels would be useless to the graphic designer.

The concept of luminance (Koenderink *et al.*, in press) works quite well in some contexts, but not at all in others. Legibility of type is almost perfectly predicted by luminance, whereas tasks involving segmentation or comparison of disjunct areas are *much* better described by equipollence. Similar observations are reported in computer vision. For instance, in cases of image segmentation, a typical Gestalt figure–ground issue, algorithms based on the simple Schopenhauer ‘parts of white’ model work much better than those based on the CIE *Lab*–distance (Sanda Mahama *et al.*, 2016). In summary, luminance fails to account for cases that depend crucially on Gestalt factors, which has obvious implications for applications in graphics and design.

If ‘luminance’ fails to predict the weight of colored patches, might it still be important in the black-and-white (B&W) case? It is this question that we address in this paper. Notice that one might well be induced to propose



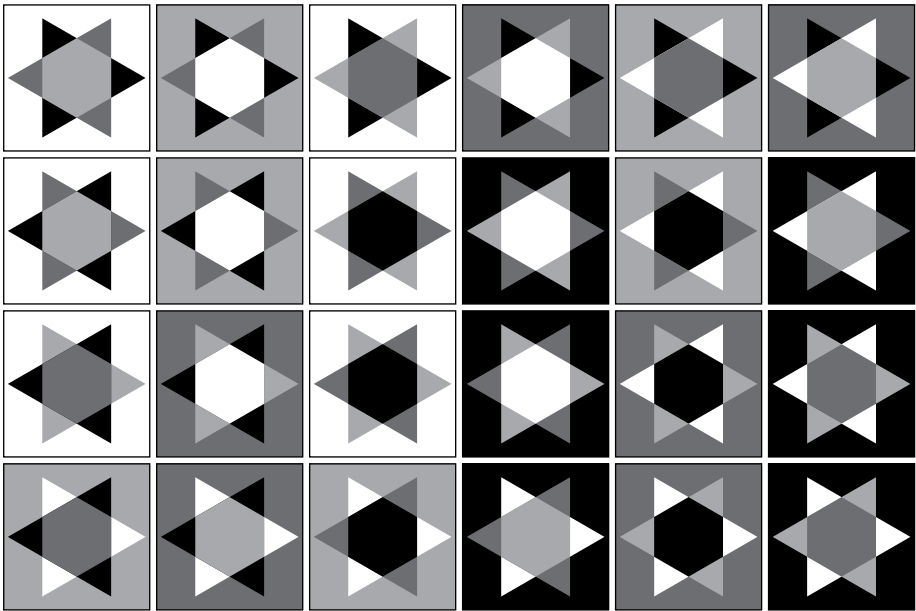
**Figure 3.** A geometrical design based on two interlocking equilateral triangles, contained in a square that acts like a frame. The intersection of the triangles generates a regular hexagon and six smaller equilateral triangles arranged in a chain. The symmetry is partly broken by grouping the small triangles in interlocking triads, thus one needs only four colors (instead of eight) to paint this configuration.

luminance as a major factor in weight. However, the example in Fig. 1 shows that the absolute boundary contrast is perhaps an even more important factor. Here we attempt to investigate the effect of luminance on weight *per se*, by minimizing the influence of boundary contrast.

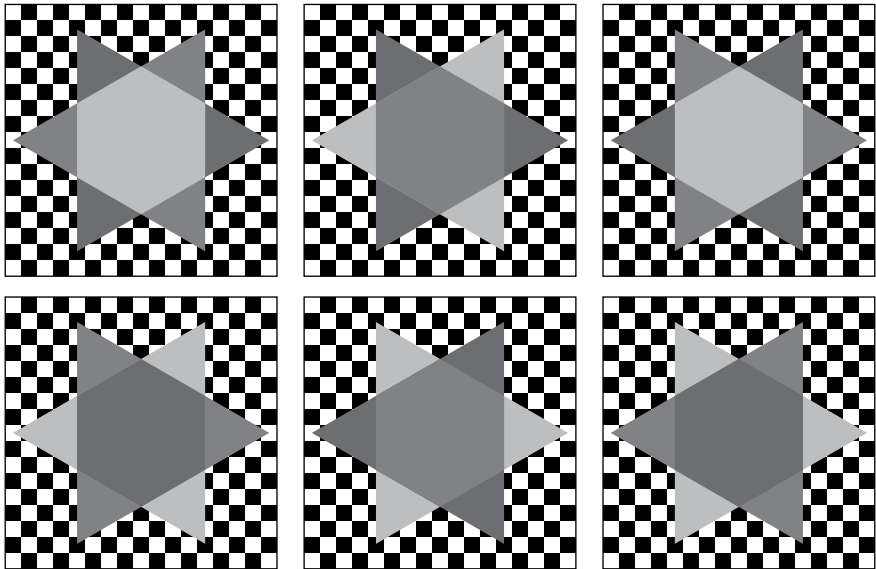
The design we use is based on the geometry shown in Fig. 3. It has a structural complexity not unlike that of a logo or simple poster design. Notice that the assignment of colors A...D is not fully general, since C and D are colors of two non-convex regions, each consisting of three disjunct triangles. This partly destroys (or rather changes) the symmetry. In a sense, the containing square is of secondary importance only, it acts like a background and frame to the whole. Many alternative designs that would be equally serviceable to this study are perfectly possible, of course. It seems a priori unlikely that they would lead to very different results, though.

In a B&W rendering one obtains various designs by assigning distinct gray tones to the areas A...D (Simmons, 1992). What counts here is the *tonal order*, rather than the photometrically precise absolute levels. Thus there exist  $4! = 24$  distinct designs. Most of these look mutually rather different, as can be seen in Fig. 4.

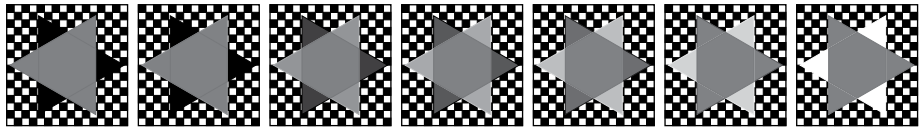
As is evident from Fig. 4, the tone of the background plays an important role. One would like to ‘take the background out of the equation’ as it were, thus simplifying the design. There are several possible methods to achieve this. For instance, painting the background pure blue works fine (see below). In Fig. 5 we render the background as *simultaneously white and black*. Although we only show a static design here, this works even better with randomly placed



**Figure 4.** Here are the 24 distinct designs based on the geometry defined in Fig. 3. One becomes aware of very distinct effects of grouping and apparent transparency. In some of these designs the two large triangles are not manifest at a cursory glance, in others they immediately strike the eye. This is of major importance in graphic design.



**Figure 5.** With the background simultaneously black and white, its role in the compositional balance disappears. Now there remain six distinct designs. Only two of these (center column) manifestly display the large triangles at cursory glance.



**Figure 6.** Here are six instances from a continuous spectrum of designs. At the left side it points to the left, at the right side it points to the right. Near the center the weights balance and the composition does not noticeably point either way. Notice that the central hexagon remains a constant, uniform mid-gray.

white and black checks that are updated at such a rate that the background appears much like a ‘snow storm’.

This removes the role of the background as a member in the sequence of tones. There remain only three tones, thus  $3! = 6$  distinct designs. They are illustrated in Fig. 5.

Only the designs in the center column of Fig. 5 clearly reveal the interlocking large triangles. These are the cases where the tone of the central hexagon is in between the tones of the two groups of small triangles. The triangles appear to be overlaid transparently (Metelli 1970, 1974).

Note that these two designs are equivalent under a mirror reflection about the vertical. The light triangle points either to the left or the right, whereas the dark triangle points in the opposite direction. (In the experiment the instruction mentions only ‘left’ or ‘right’, but – of course – it is also possible to ‘see’ oblique pointing directions. We ignore these here.) In the examples of Fig. 5 center column, the pointings of the light and dark triangles (at least roughly) balance each other. By adjusting the tones appropriately, thus changing the ‘weights’ of the light and dark triangles, one may make the design *as a whole* point left or right. This is an example of the compositorial balance discussed above, with the advantage that the balance is easily parameterized (Fig. 6).

## 2. Methods

### 2.1. Display

The display is the screen of an Apple MacBook Pro 15” (mid-2007 model). It was spectrophotometrically calibrated with a X-Rite ColorMunki, using the Argyll software. Luminance and uniformity were double checked with an Asahi spotmeter. The luminance of the display white was  $402 \text{ cd/m}^2$ . The display was linearized using the Bergdesign SuperCal method (To *et al.*, 2013). The white point was  $x = 0.319$ ,  $y = 0.348$ .

We believe the precise calibration to be largely irrelevant. We merely report it here in order for the experiment to be replicated precisely.

## 2.2. Presentation

Observers viewed the display binocularly from a distance of 57 cm, using their preferred optical correction when necessary. The full screen measured  $20^\circ \times 33^\circ$ .

The presentation software was written in Processing2+, a variety of Java aimed at artists and designers. It allows the full spatial, temporal and colorimetric control for our needs and allows fast development. User interaction was limited to the use of the left–right and up–down arrow keys, whereas the space bar signified that the user initiated the next trial. We provided no fixation mark, viewing was free.

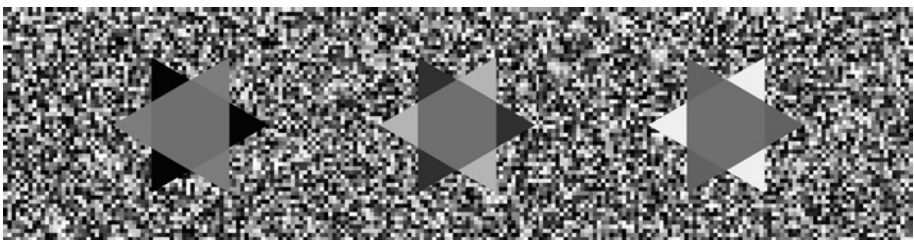
## 2.3. Observers

Seventeen naive observers participated in the experiment. Observers were predominantly female.

## 2.4. Experiment

Gestalt factors crucially depend on the balance of colors (Takahashi *et al*, 2010). In some cases (Figs 5 through 7) one becomes immediately aware of an arrow pointing left, in other cases of an arrow pointing right. It depends upon the colors and/or tones of the parts. For some choices it becomes very hard to decide whether the configuration points to the left or to the right (Figs 6 and 7 center). This may be taken as an operational definition of ‘equal weight’.

A problem is how to pick a suitable ‘neutral’ (neither white nor black, nor anything in between) background. We used two implementations of the paradigm. In method A the background is blue, in method B it is a spatiotemporal ‘snowstorm’ pattern (Fig. 7). The latter ‘has no tone’, the former has no well



**Figure 7.** The principle of the ‘monochrome Gestalt paradigm’, this is a momentary impression of method B. It can only be suggested rather than illustrated, in reality the noise was much finer and was refreshed at frame-rate. The attempt was to create a background that would not fit anywhere on the tonal range from black to white. This succeeds to the effect that the snowstorm background contrasts about equally with any uniform patch no matter the tone, from white to black.

defined tone either, for although the luminance functional suggests it has an almost black tone, the blue actually looks bright. Moreover, any gray tone contrasts with the blue. We consider both backgrounds to be ‘neutral’ with respect to tone, in the sense that they cannot sensibly be assigned any tone.

There are three distinct regions (apart from the background), namely two composed of three triangles each (areas C, D in Fig. 3) and a central hexagon (area B in Fig. 3). The central hexagon has a gray tone  $1/2$  (on a 0–1 scale), the two other regions have gray tones  $(1+p)/4$  and  $1/2+(1+p)/4$ , where the parameter  $p$ , controlled by the observer, varies between the limits  $-1$  and  $+1$ .

In each case a warning signal (far outside the stimulus area) is provided to notice the observer that one or the other limit of the parameter range ( $\pm 1$ ) has been reached.

Observers were instructed to start a trial by first looking at either extreme. Then, starting from a random value, to use large increments to find the approximate environment of the point of balance, and finally to switch to small increments to do their setting. In case the region of subjective equality is ill defined, all the observer can do is indicate the approximate ratio using the large increments. Otherwise a single small increment might be close to a just noticeable difference.

In each presentation participants did 10 settings, each time starting from a fresh random value of  $p$ . Half of the observers did the experiments in the sequence AB, the other half in the sequence BA. We embedded this experiment in another, larger one (Koenderink *et al.*, in press), thus the A and B presentations were separated by other (similar, but chromatic) experiments, ensuring that they may be considered mutually independent.

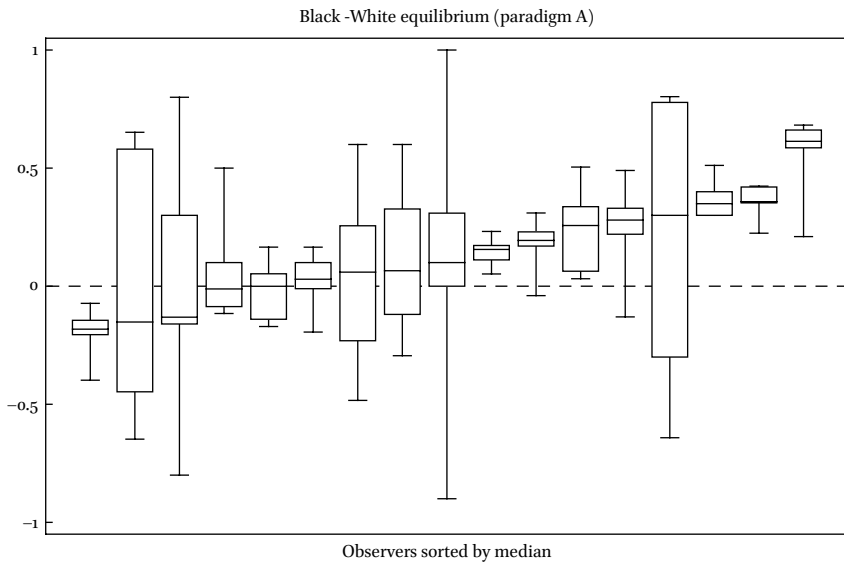
Most observers immediately recognize that some ‘equilibrium’ *has* to exist between the limits, so much is evident from even a cursory look at the limits. However, some find it hard to indicate the exact location of an equilibrium point. Results for all observers are shown in Figs 8 and 9. For the sake of clarity, we have sorted the observers by the median of their response in these figures.

### 3. Analysis

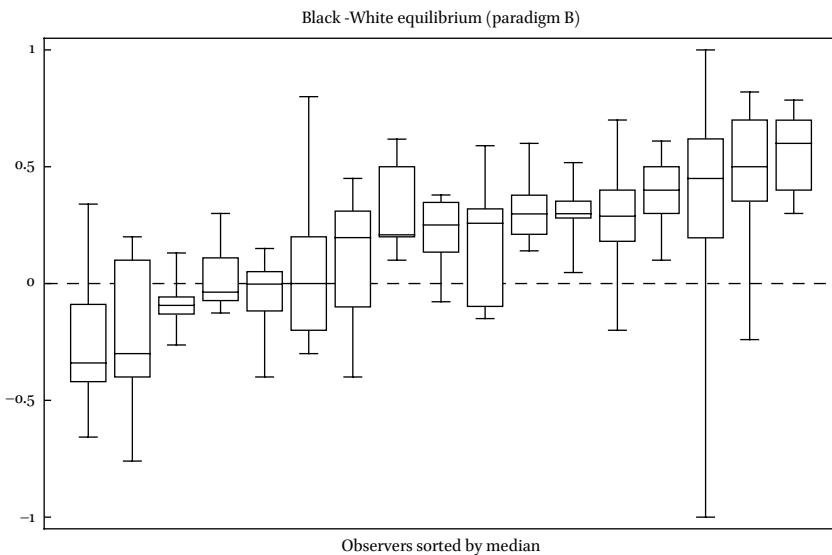
What to make of these results? Evidently, participants can do the task and consider it a reasonable one. Most observers arrive at a position midway between black and white, some with apparent precision whereas others spread their estimates over a large part of the range. There is an evident asymmetry between the light and dark directions, possibly due to our particular choice of the ‘mid-gray’.

There is a significant spread of settings, some being at the negative, others at the positive side of the zero level. The interquartile ranges for the observers are also very different.

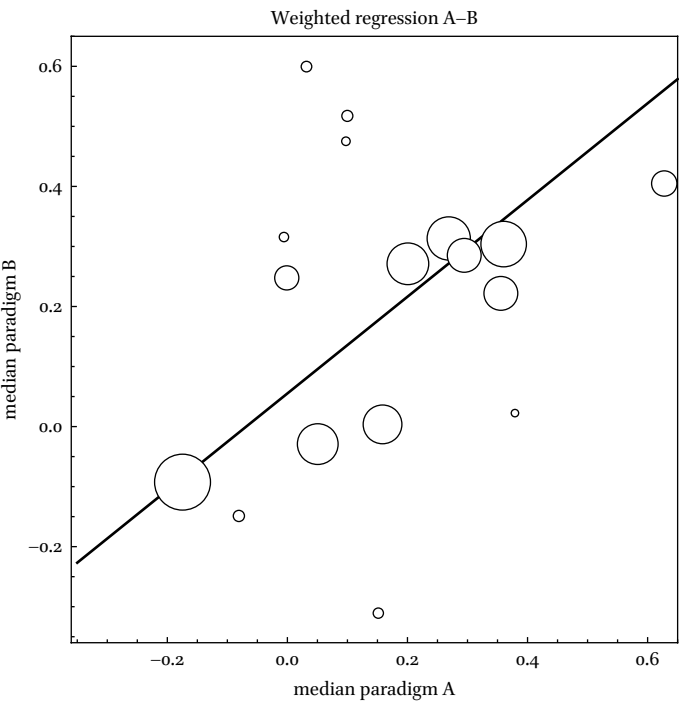




**Figure 8.** Settings for paradigm A (blue background) for all observers. Each observer did 10 repeats. The box-and-whisker plots show total range and quartiles. A positive value corresponds to the right side, a negative value to the left side of the stimuli shown in Fig. 7 (apart from the background).



**Figure 9.** Settings for paradigm B ('snow storm') for all observers. Each observer did 10 repeats. Like in figure 8, the box-and-whisker plots show total range and quartiles. A positive value corresponds to the right side, a negative value to the left side of the stimuli shown in Fig. 7.



**Figure 10.** Weighted regression between the medians for paradigms A and B over all participants. Black line shows the fit. The diameters of the dots show the weights, which are taken to be inversely proportional to the interquartile ranges.

Are the settings—say the medians—random or idiosyncratic? One way to obtain an idea is to compare the observer’s responses for the two tasks. This will obviously not be informative for the observers with very high spreads, though. It might be informative for those observers who have a narrow response range. The medians of the two cases are not strongly correlated (Kendall’s tau 0.18), whereas the interquartile ranges for the two cases are positively correlated (Kendall’s tau 0.4), suggesting that the results are at least to some extent idiosyncratic.

A weighted Pearson product-moment correlation coefficient for the medians (Wolberg, 2005), using the reciprocals of the interquartile ranges as weights, yields 0.65. Figure 10 shows a scatterplot—including weights—of the regression. It may be concluded that the ‘focussed’ observers, those with small spread, show marked idiosyncratic differences, whereas the others are essentially at random: an unweighted Pearson product-moment correlation yields a coefficient of only 0.28.

Apparently, the differences between methods A and B are only slight, somewhat strengthening our conviction—based on pure phenomenology—that the

background is hardly a factor in either case. However, there certainly exist differences.

The extremely large interquartile ranges are limited to case A. It involves only two or three observers, but in view of the fact that these same observers were in the normal range in case B, suggests that we spot an idiosyncrasy here. Unfortunately, we can hardly make a point of that, since one would need at least an order of magnitude increase in the number of observers to address such an issue seriously.

The weighted regression reveals an offset of 0.06, which is hardly informative, and a slope of 0.81. Thus perhaps the medians are somewhat tighter in case B, but the difference is small enough that we do not want to make a point of it.

The overall result is that

- the participants react differently in a quantitative, though not a qualitative, sense;
- cases A and B lead to essentially equivalent results;
- the median settings range over 0.80 (case A) or 0.94 (case B), which amounts to almost half the full range (–1 to +1, thus a range of 2.0).

The main impression is that the idiosyncratic differences are perhaps surprisingly large. They are certainly much more impressive than what we found for chromatic cases (Koenderink *et al.*, in press). The fact that participants agree qualitatively may be taken as an indication that the *order* of luminances is much more important than the absolute magnitudes of the luminances per se.

The latter interpretation agrees with the informal observation that the transfer function from nominal luminance to display gray level is largely irrelevant as long as it is monotonic. For the graphical arts this is a crucial fact, because the actual transfer functions differ widely for the various reproduction processes. Even a cursory investigation of the various renderings of some given work on the Internet will soon drive that point home.

#### 4. Conclusions

Although the B&W paradigms clearly work, since they are accepted by most naive observers as reasonable tasks, fact is that the results are qualitatively similar, but quantitatively rather variable. That there exist such significant interindividual differences seems interesting and possibly of importance in the graphical arts.

It seems likely that the effect of compositions depending on Gestalt factors playing between areas of different gray tones may give rise to rather different feelings of ‘balance’. The stimulus patterns A, B are already in the ballpark

of structural complexity of many strong logo or poster designs, so the findings reported here might well apply to such cases. Even small differences may give rise to semantically meaningful effects, for instance, in the present case, whether the figure ‘points left’ or ‘points right’.

As a matter of course, such paradigms, whether chromatic (Koenderink *et al.*, in press) or monochrome (B&W, present study), are sensitive—sometimes even *very* sensitive in the sense that a minor change may completely upset the composition—to the particular parameter choices. These choices do not only involve tone, but equally involve location in the frame, size and shape, edge quality and so forth.

Apparently a wide area of investigation lies open here. It evidently relates marginally to aesthetic judgments, although observers are much more certain about their judgments in the present case. It also relates marginally to classical visual psychophysics, although the degree of idiosyncrasy is much higher and the relation to the physical stimulus structure much weaker. The judgment of weight and balance is a rewarding realm of experimental phenomenology that holds a key position (perhaps Rosetta stone-like) between the pure phenomenology of aesthetics and the barren thresholds of psychophysics.

## References

- Buck, S. L. (2014). Dark versus bright equilibrium hues: rod and cone biases, *J. Opt. Soc. Am. A* **31**, 75–81.
- Commission Internationale de l’éclairage (CIE) (1990). CIE 1988, 2° spectral luminous efficiency function for photopic vision. Publ. No. 86, Bureau Central CIE, Paris, France, pp. 1–11.
- Dondis, D. A. (1973). *A Primer of Visual Literacy*. MIT Press, Cambridge, MA, USA.
- Eisner, A. and MacLeod, D. I. A. (1980). Blue-sensitive cones do not contribute to luminance, *J. Opt. Soc. Am. A* **70**, 121–123.
- Graves, M. (1951). *The Art of Color and Design*. McGrawHill, New York, NY, USA.
- Koenderink, J. J., Van Doorn, A. J. and Gegenfurtner, K. (in press). Color weight photometry, *Vis. Res.* doi:10.1016/j.visres.2017.06.006.
- Locher, P., Overbeeke, K. and Stappers, P. J. (2005). Spatial balance of color triads in the abstract art of Piet Mondrian, *Perception* **34**, 169–189.
- MacManus, I. C., Edmondson, D. and Rodger, J. (1985). Balance in pictures, *Br. J. Psychol.* **76**, 311–324.
- Metelli, F. (1970). *An Algebraic Development of the Theory of Perceptual Transparency. Contemporary Problems in Perception*. Taylor and Francis, London, UK.
- Metelli, F. (1974). The perception of transparency, *Sci. Am.* **230**, 91–98.
- Mokaran, M. A. (2007). *Visual Balance in Engineering Design for Aesthetic Value*. Thesis, University of Saskatchewan, Saskatoon, SK, Canada.
- Monroe, M. (1926). The apparent weight of color and correlated phenomena, *Am. J. Psychol.* **36**, 192–206.

- Morriss, R. H., Dunlap, W. P. and Hammond, S. E. (1982). Influence of chroma on spatial balance of complementary hues, *Am. J. Psychol.* **95**, 323–332.
- Parada-Castellano, R. (2016). Study of balance of images using visual weight, *COLOR Res. Appl.* **41**, 175–187.
- Pinkerton, E. and Humphrey, N. (1974). The apparent heaviness of colours, *Nature* **250**(5642), 164–165.
- Sanda Mahama, A. T., Dossa, A. S. and Gouton, P. (2016). Choice of distance metrics for RGB color image analysis, *IS&T International Symposium on Electronic Imaging 2016, Color Imaging XXI: Displaying, Processing, Hardcopy, and Applications*, San Francisco, CA, USA, pp. 349.1–349.4.
- Schopenhauer, A. (1870). *Über das Sehnen und die Farben*, 3rd ed. Julius Brockhaus, Leipzig, Germany.
- Simmons, G. (1992). *The Technical Pen, Techniques for Artists*. Watson-Guptill, New York, NY, USA.
- Smith, V. C. and Pokorny, J. (1987). Is there a luminance channel? *Farbe* **34**, 123–128.
- Speed, H. (1913). *The Practice and Science of Drawing*. Seeley, Service & Co., London, UK.
- Takahashi, S., Ohya, K., Arakawa, K. and Ishisaka, Y. (2010). Perceived strength of edge, depth and brightness of the Kanizsa illusion as a function of the color contrast between figures and background. *Gestalt Theory* **32**, 155–166.
- Taylor, J. F. A. (1964). *Design and Expression in the Visual Arts*. Dover, New York, NY, USA.
- To, L., Woods, R. L., Goldstein, R. B. and Peli, E. (2013). Psychophysical contrast calibration, *Vis. Res.* **90**, 15–24.
- Wise, B. K. and Wise J. A. (1988). *The Human Factors of Color in Environmental Design: A Critical Review*. NASA contractor report 177498. Retrieved from <http://spacearchitect.org/pubs/NASA-CR-177498.pdf>; last accessed 3 July 2017.
- Wolberg, J. (2005). *Data Analysis Using the Method of Least Squares: Extracting the Most Information from Experiments*. Springer, Berlin, Germany.
- Wong, W. (1993). *Principles of Form and Design*. Van Nostrand Reinhold, New York, NY, USA.
- Wright, B. (1962). The influence of hue, lightness and saturation on apparent warmth and weight, *Am. J. Psychol.* **75**, 232–241.