Pictorial balance is a bottom-up aesthetic property mediated by eye movements

David Corwin
Dcorwinmd[@]gmail.com

Abstract

This paper identifies a specific pattern of luminance in pictures that creates a low level non-subjective neuro-aesthetic effect. Pictures evoke both a top down and a bottom-up visual percept of balance. Through its effect on eye movements, balance is a bottom-up conveyor of aesthetic feelings of unity and harmony in pictures. Eye movements are predominantly influenced by the large effects of saliency and top-down priorities so that it is difficult to separate out the much smaller effect of balance. Given that balance is associated with a unified and harmonious picture and that there is a pictorial effect known to painters and historically documented that does just that, it was thought that such pictures are perfectly balanced. Computer models of these pictures were created by the author on LCD monitors. They were found to have bilateral quadrant luminance symmetry with a lower half lighter by a factor of ~1.07 +/- ~0.03. A top weighted center of quadrant luminance calculation is proposed to measure balance.

To show that this effect exists, two studies were done that compared identical pictures in two different frames with respect to whether they appeared different given that the sole difference is balance. A conventional study of preference could not be done because LCD images use polarized light interfering with luminance contrast perception. This creates an effect similar to glare diminishing the aesthetic quality of the effect while the effect on eye movement is still present. Results show that with observers, mostly painters, there was a significant correlation between average pair imbalance and observations that two identical pictures appeared different. This indicated at a minimum that the equation for calculating balance was correct.

The effect is the result of the absence of forces on eye movements created by imbalance. With painters who can disregard salience, this is manifested by the absence of forces drawing attention to any part of the image. All parts including that in peripheral vision receive attention, and the eye seems to slide through rather than to jump from objet to object. This effect is called pictorial coherency. Large tonally contrasting forms, geometric forms or many different forms that cannot be visually combined prevent it from being seen. This unaccustomed force of imbalance explains why viewing pictures cause fatigue to anyone who does not constantly study pictures. That pictures can evoke a low level percept of balance would indicate that it belongs to earlier evolutionary
development that the author believes is used to follow movement by defining a complex object as a simple vector.

Introduction

A picture, a bounded flat image, corresponds to nothing found in nature or mental imagery. This includes the borderless images in prehistoric caves [1]. It differs from an image without borders in that it evokes the percept of balance: the parts relate to both the center and the borders [2]. Balance is both a bottom-up percept and a top-down aesthetic concept taught in schools. But balance of what and how? Color, form or tonal areas have been proposed. These somewhat immeasurable qualities have been said to be balanced around a geometric center or the principal pictorial axes and defined by a center-of-gravity type of equation. This is said to explain both the perception of balance and the aesthetic qualities of unity and harmony in pictures [2,3,4]. It is either impossible or would require too much calculation for the visual system to do this. Any attempt to show that low level balance is determined by some sort of equation might be associated with this erroneous concept. However, in the same way that elements of salience such as color, contrast, and orientation are used to calculate eye movements within a picture, one might be able to calculate balance as a function of pictorial properties [5].

Eye movements are predominantly influenced by saliency and top-down priorities [5,6,7]. These determinants of eye movements are large so that it is difficult to distinguish the much smaller effect of balance. Locher and Nodine tried to show the effect of balance by comparing eye movements in pictures that were thought to be balanced with variations of these pictures. Using trained and untrained observers, they found that trained observers had more diverse and fewer specific exploratory movements in the more balanced compositions [8,9,10]. These effects are characteristic of better balance in that they indicate that the trained eye is less constrained to specific salient areas but do not quantify the effect of imbalance.

Given that the determination of how the visual system calculates pictorial balance can only be done empirically, I thought that one approach to the problem might be to create a perfectly balanced picture. Such a picture where balance has no effect on eye movements could be used as a base from which other measurements would be made. Perfect balance has also been used to explain the feelings of unity and harmony evoked by some pictures, and this corresponds to an obscure effect observed by painters that does just that. There is no name or metaphorical description for the percept evoked by these pictures other than to say that they exhibit the aforementioned effect. It is only discussed when before such a picture in art schools. The unexpected change in a painting from being remarkably unified and harmonious to something less so or the reverse is quite striking. Peripheral
vision seems to become more salient, creating the feeling that the whole picture can be seen at one time without need to focus on any pictorial element. I call this percept pictorial coherence.

This was first noted by Roger de Piles, a French critic and painter around 1700 who described an arrangement of “lights and darks” in paintings creating an effect that is spontaneous and intense giving the impression that the viewer could gaze at the entire painting without focusing on any particular form [11]. Cézanne apparently noted this. Indeed, his inability to describe this effect gives some of his conversations a cryptic quality [12, p.88-89, 110-111]. Both Delacroix 1923 and Kandinsky give vivid descriptions of seeing it in paintings [13,14,15]. (addendum A)

To study this effect I created computer models of images that evoke it through experience acquired from many years of painting. Image analysis indicated that the percept was obtained when an LCD picture, centered at eye level and parallel to the plane of vision, has bilateral quadrant luminance symmetry with the lower half having slightly less luminance than the upper half by a factor of ~1.07 ± ~0.03 (where luminance is measured from 1 to 255 with black equal to 1 as measured by Photoshop CR6). A theoretical formula for balance can be derived from this which defines the observed state of perfect balance and explains observations implying a center-of-mass-like effect: if a rectangular picture with bilateral quadrant luminance symmetry and a darker upper half can be said to be perfectly balanced i.e. its center of luminance is located at the geometric center of the picture, then the standard formula for the center of mass of four connected objects can be used. In this formula quadrant luminance \( L_{xQ} \) replaces the mass of each object with the picture centered at \((0,0)\), the geometric center of each quadrant is its center of mass, and \( L_{TOTAL} \) is taken as the average total luminance.

\[
\begin{align*}
X &= \left( X_{ULQ}L_{ULQ} + X_{URQ}L_{URQ} + X_{LLQ}L_{LLQ} + X_{LRQ}L_{LRQ} \right) / L_{TOTAL} \\
Y &= (1.07)Y_{ULQ}L_{ULQ} + (1.07)Y_{URQ}L_{URQ} + Y_{LLQ}L_{LLQ} + Y_{LRQ}L_{LRQ} )/L_{TOTAL} \\
L_{TOTAL} &= (L_{ULQ} + L_{URQ} + L_{LLQ} + L_{LRQ})/4, \\
\text{Balance} &= \sqrt{X^2 + Y^2} \text{ or a vector } L^* = (X,Y)
\end{align*}
\]

Y values of the upper quadrants are modified by 1.07 to account for the increased visual weight of the upper half. \( X_{ABQ} \) and \( Y_{ABQ} \) are the coordinates of the center of their respective quadrants, and \( L_{ABQ} \) is a quadrant’s average luminance.

While the aesthetic effect can be intense with a picture seen by reflected light, the percept is much less so in LCD pictures. There is a diminished feeling of unity while the effect on
eye movements is maintained. Because this is more subtle, a study of subject preference could not be done. Bertamini et al. [16] noted that observers preferred a print reproduction to seeing the same picture in a mirror or on a monitor and that they preferred the mirror view of the actual image to the image on an LCD screen. Locher P, Smith JK, et al. [17] made similar observations. Light from LCD screens is polarized to which humans are sensitive. Misson and Anderson [18] showed that this polarization interferes with luminance contrast perception. This increases salience and explains the difference between and LCD image and a reflected light image. LCD images had to be used to permit precise luminance determination and reproducibility. Using OLED images but not AMOLED would be preferable if the study is repeated to eliminate the effect of polarization.

With respect to the effect an image framed with a white border against a black ground is seen as one visual object – the image and the frame. This is not true for a picture framed in black against a white ground. The former permits the comparison of a balanced framed picture with the identical picture in a slightly different frame that renders the combination unbalanced. This avoids the confounding variable of salience that occurs when an image is compared with even a small variant. Unfortunately, the white border also increases the overall saliency of the picture. If an observer is able to compare the two images and to disregard the frame, any perceived difference would be ascribed to the effect. Most people, who observed the two identical pictures in succession, saw no difference. Only a group of painters and a few people very interested in looking at pictures were found likely to see it. Two studies were done to show that those likely to see the effect in paintings could notice a difference between “identical pairs.” A painting exhibiting coherence is said to be perfectly balanced; all others will be called unbalanced.

**Materials and Methods**

Each study consisted of ten pairs of images: Five in which one of the images was coherent and the other unbalanced (balanced or coherent pairs) and five in which both were unbalanced (unbalanced pairs). Images were prepared using Photoshop©. Quadrant image luminance characteristics can be changed in an unobtrusive and precise manner in Photoshop through the use of the level command on a small irregular form in which the luminance is usually changed by no more than ±5 percent. The form is then moved around until the desired quadrant value is reached. The two studies differed only in using different sets of white borders. The three possible image and border combinations are in Fig. 1. Study I compared figure 1a with either 1b or 1c. Study 2 compared figure 1b with 1c. Only the white borders are seen when a picture is seen against a black ground, and in the study the viewer sees only variations in the borders. Observers see sequentially a picture with one frame followed by the other on a calibrated iPad using the ColorTrue™ app and a Colorite™ device. The observers viewed the iPad centered and parallel to the
plane of vision at arm's length. The images on the iPad were approximately 5x6.75 inches and subtended visual angles of roughly 19° and 23° respectively under the conditions of the study. The pictures used for the studies are in Fig. 2. Images 1, 3, 5, 7, 8 were the balanced pairs (all images are in the supplemental files).

Fig. 1 The three possible image and border combinations

Fig. 2 Pictures used in the study pairs
The observers were predominantly artists of any age who gave their consent orally to participate in an anonymous study about how people look at pictures. They were told that it consisted of looking carefully at ten pairs of pictures and being asked whether the two central images appeared to be the same or different. They were also told that it was thought that some pairs might be seen as the same and others as different. No identifying information or medical history was obtained; there was no attempt to eliminate observers who were color deficient, had cataracts or had impaired binocular vision. The research study was done in accordance with the Helsinki Declaration (WHO). 45 observers were included in the first study and 39 observers in the second study. Other observers were excluded if they could not follow the directions and either rapidly said without careful examination that they could see no difference or if they insisted that the frame change.
was the only difference. There were only four of the latter. For pairs in which both pictures are unbalanced, it was presumed that observers would find the pictures to be identical so that the response of “same” was considered correct. With the balanced pairs a response of “different” would be labeled correct.

Observers viewed the pictures sequentially on an iPad at arm’s length with the examiner seated slightly behind them. They were permitted, indeed encouraged, to hold the iPad themselves as long as it was maintained correctly centered and parallel to the plane of vision. There were no restrictions on the length of observation, and observers could return to a previous image as much as they wanted. However, subjects were told that the difference if any was more in the way of a feeling than a distinct difference and were discouraged from making point by point comparisons. The difference was described as analogous to that between monophonic and stereophonic music: same music but seems different.

**Results and discussion**

Four observers could identify all the balanced pairs correctly, and 5 observers made one error. (Table 1) Although it cannot be statistically proven, it is very suggestive that the effect is being seen. Some subjects thought they saw differences in color, one thought the depth of field was greater, and many saw differences but could not describe them.
Among observers who perceived differences in any pair, many saw differences in both balanced and unbalanced picture pairs. Initially this was thought to be due to guessing. However, there is a correlation between the percent of pairs observed to be seen as different and the pair average distance (average imbalance). Pairs can be divided into two groups: one in which the average state of imbalance within the pair is small (the balanced pairs), and the other where the average imbalance is large. The balanced pairs were identified correctly by 40% of observers, and the unbalanced pairs were identified correctly by 75% of the observers. The correlation between average pair distance and the percentage identified as different was: $r_{18} = -0.723$, $p < 0.001$. (Table 2).

This might verify the hypothesis that observers see pairs as different within balanced pairs. However, it might just indicate that observers see pairs that have low average pair imbalance as different from those having high average pair imbalance. One could verify this by constructing unbalanced pairs with an average imbalance no different from that found in coherent pairs.

If one were to select unbalanced pairs that contain particular types of salient features, then the results would be similar to image pairs 2 and 10 which most subjects saw as the same. These pictures are not only unbalanced but contain salient images such as the woman in white that strongly draw one’s attention. Other salient types of forms or groups of forms will be discussed below. Therefore although table 1 might appear to show a Gaussian distribution, the answers are not a simple question of chance. It is like the
aforementioned analogy with monophonic and stereo music; one can place two speakers so far apart that anyone can identify a stereo effect or move them so close that no one can. One can liken the image pairs to pairs of flags extending in the distance until they appear to everyone to be one flag. Using this analogy one could say that this particular study does not distinguish between the people who see two flags and the few who become fascinated by what is painted on some. The results indicate at a minimum that the visual system sees pictures in terms of the formula of quadrant luminance balance or a variation of it that corresponds to distance in the analogy. Additional image pairs can be found in the supplemental files.

<table>
<thead>
<tr>
<th>image pair</th>
<th>average pair distance</th>
<th>% correct/different</th>
<th>image pair</th>
<th>average pair distance</th>
<th>% different</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st Study</td>
<td></td>
<td></td>
<td>1</td>
<td>0.462</td>
<td>0.489</td>
</tr>
<tr>
<td>3</td>
<td>0.318</td>
<td>0.356</td>
<td>4</td>
<td>1.894</td>
<td>0.289</td>
</tr>
<tr>
<td>5</td>
<td>0.427</td>
<td>0.444</td>
<td>6</td>
<td>1.679</td>
<td>0.333</td>
</tr>
<tr>
<td>7</td>
<td>0.509</td>
<td>0.378</td>
<td>9</td>
<td>0.944</td>
<td>0.178</td>
</tr>
<tr>
<td>8</td>
<td>0.522</td>
<td>0.444</td>
<td>10</td>
<td>2.197</td>
<td>0.111</td>
</tr>
<tr>
<td>2nd Study</td>
<td></td>
<td></td>
<td>1</td>
<td>0.608</td>
<td>0.410</td>
</tr>
<tr>
<td>3</td>
<td>0.652</td>
<td>0.359</td>
<td>4</td>
<td>2.153</td>
<td>0.154</td>
</tr>
<tr>
<td>5</td>
<td>0.473</td>
<td>0.436</td>
<td>6</td>
<td>1.934</td>
<td>0.385</td>
</tr>
<tr>
<td>7</td>
<td>0.658</td>
<td>0.282</td>
<td>9</td>
<td>1.389</td>
<td>0.256</td>
</tr>
<tr>
<td>8</td>
<td>1.026</td>
<td>0.436</td>
<td>10</td>
<td>2.455</td>
<td>0.128</td>
</tr>
</tbody>
</table>

The use of an iPad in different lighting environments was necessary to bring the study to the observers as no laboratory was available. Reflections from the iPad screen increased the number of false negatives, i.e. coherent pictures being seen as unbalanced. The possibility of examiner influence cannot be excluded although an attempt was made to diminish this by being seated behind the subjects. The results of the unbalanced pairs were contrary to the expectation that no one would see the unbalanced pairs as different as seen in table 2. After all they are identical pictures. A MacNemar test showed data validity $p < 0.01$. 

Table 2 Pair average distance and percent of observers that identified the pairs as different.
Pictorial balance is extremely sensitive to small differences of quadrant luminance which explains its fugacity without precise viewing conditions (table 3). Changing temperature of the lighting or height at which the painting is hung with respect to the viewer destroys the effect. With pictures in a studio even the drying of paint in the following twelve hours creates a tonal change that is sufficiently to make a coherent state disappear. Given that not everyone can see the effect, that it is normally arrived at through chance and is usually fleeting, and that it has no name would explain why it is not discussed in the 20th century literature.

Table 3 shows the lateral balance and top/down ratios for the 2nd study. The unbalanced pairs are in bold red.

<table>
<thead>
<tr>
<th>Pair</th>
<th>lateral balance</th>
<th>top/down</th>
<th>pair</th>
<th>lateral balance</th>
<th>top/down</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.05</td>
<td>0.97</td>
<td>6</td>
<td>1.04</td>
<td>1.37</td>
</tr>
<tr>
<td>2</td>
<td>1.03</td>
<td>0.95</td>
<td>7</td>
<td>0.94</td>
<td>0.97</td>
</tr>
<tr>
<td></td>
<td>0.99</td>
<td>1.01</td>
<td>1</td>
<td>1.07</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>1.06</td>
<td>8</td>
<td>1</td>
<td>1.08</td>
</tr>
<tr>
<td></td>
<td>0.93</td>
<td>1.16</td>
<td>1</td>
<td>1.26</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>0.98</td>
<td>0.9</td>
<td>9</td>
<td>1.05</td>
<td>0.87</td>
</tr>
<tr>
<td></td>
<td>1.06</td>
<td>0.83</td>
<td>1</td>
<td>0.98</td>
<td>1.01</td>
</tr>
<tr>
<td>5</td>
<td>0.95</td>
<td>1</td>
<td>10</td>
<td>0.8</td>
<td>0.96</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>1.07</td>
<td></td>
<td>0.86</td>
<td>1.06</td>
</tr>
</tbody>
</table>

Conclusions

When an innately sensitive observer views a picture ignoring saliency, balance remains to direct the eye. With perfect pictorial balance there is no controlling effect on eye movements. One becomes attentive to the whole picture and is able to move smoothly through it instead of being attracted to one form after another. This is what Roger de Piles describes as seeing the whole picture at one time “le tout ensemble” [14] 2000). It was found empirically that artists particularly painters had to be used as observers. This corresponds to results showing that painters view pictures differently than untrained individuals [10, 19,20,21,22]. Antes showed that fixation patterns distinguish painters from non-painters, and Koide et al [23] using a prediction model of a saliency map.
showed that artists are less guided by local saliency than non-artists. Coherence facilitates this in an extraordinary way.

Geometric, strongly contrasting or many different equally prominent forms that cannot be combined conceptually force the eye to focus on each particular form. While these pictures may meet the luminance balance criteria, they will probably not be seen or are much less likely to be seen as coherent. They prevent the viewer from disregarding salience and force the eye to go from salient feature to salient feature. One could say that total pictorial imbalance is not simply a calculation of luminance balance but also a measure of the degree of formal saliency.

Normal viewing seems effortless because we are accustomed to the work performed by the eye muscles. However, when we look at pictures carefully, the pull of balance creates unaccustomed work which is experienced as fatigue. No one complains of visual fatigue when intently watching a moving image on an LCD screen where balance plays no role, but they do when looking carefully even briefly at pictures. A study of how long subjects gaze spontaneously at paintings of similar subjects without particularly salient forms but with differing degrees of imbalance might show this. Like any training, repeated exposure diminishes this fatigue. This could be done with OLED monitors (but not AMOLED) because the light is not polarized.

A flat surface with no marks, i.e. one coated with a thick uniform layer of paint but modified so that the upper half is slightly darker, also cannot be seen or perhaps very weakly felt as coherent. There has to be some surface quality for this part of the visual system to recognize an object. It has been shown that forms in nature have a fractal quality, that fractal images have an aesthetic quality and that the visual system has evolved to respond to natural conditions [24,25,26]. Therefore, it can be inferred that the part of the visual system that calculates balance is also most sensitive to fractal forms. The author has observed that binocular vision is necessary to observe the percept.

The explanation of the particular phenomena of coherence as a result of eye movement is only a partial explanation. Perhaps its most marked characteristic is that the attention is riveted in an indescribable way. It is, as Delacroix describes it so eloquently, different from the peak effects of other art forms - no tearing, shivering or trembling [27,28,29] so that its neural origins should be searched elsewhere.

Pictures are a very recent invention requiring technology to create flat surfaces (maybe 4th or 5th BC). They are not part of the evolutionary development of humans. Borders change an image into a picture, a visual object evoking the percept of balance. Since balance is a low level percept resulting from object luminance, it must have evolved for
some other reason. It is doubtful that something as complex as pictorial balance is an accident without evolutionary significance such as has been suggested for the visual perception of polarized light [18]. I suggest that it is a very primitive mode of identifying objects and their movement as it can represent a complex object as a moving vector.

One can object to the results of this paper as dehumanizing the art of painting as if all one has to do is paint by numbers. This might have been the response of a teacher of painting who has an excellent eye but who absolutely refused to cooperate once I explained the research. This is perhaps a legitimate but limited critique. Low level balance simply takes a composition that may or may not be particularly poetic (substitute whatever term you prefer) and at least makes it easier to look at and at best under the right conditions makes it stunning. When I saw Vermeer’s *The Art of Painting* in Vienna, it was exhibited under conditions at that time that made it completely unbalanced and difficult to look at. I did not have this response at all when I saw it in Washington at an exhibition at the National Gallery. As music’s effect can depend on the performance, so a painting’s effect depends on how it is hung. Two musicians play the same music but one performance is most moving while the other is not at all so. What’s the difference? Finding the quantitative difference in the qualitative evaluation of performance or depiction is the most important aspect of studying the neuro-aesthetic response. It is the understanding of the subtle distinction in neural responses between “okay” and “great” for a subjective experience, and then using this to help the artist.

It is impossible to have this paper published in a reputable journal because an institutional review board approval is required for human studies, and the author is not a member of any institution. Correspondence is welcome.

**Addendum A**

Eugene Delacroix:

"There is a kind of emotion which is entirely particular to painting: nothing [in a work of literature] gives an idea of it. There is an impression which results from the arrangement of colours, of lights, of shadows, etc. This is what one might call the music of the picture…you find yourself at too great a distance from it to know what it represents; and you are often caught by this magic accord. In this lies the true superiority of painting over the other arts [literature], for this emotion addresses itself to the innermost part of the soul…[and] like a powerful magician, takes you on its wings and carries you away.” [13,14]
Wassily Kandinsky describes it, but without being aware of de Piles’ observations which would have permitted him to understand his sensations, drew the conclusion that it resulted from the absence of recognizable objects:

“It was the hour of approaching dusk. I returned home … when suddenly I saw an indescribably beautiful picture, imbibed by an inner glow. First I hesitated, then I quickly approached this mysterious picture, on which I saw nothing but shapes and colors, and the contents of which I could not understand. I immediately found the key to the puzzle: it was a picture painted by me, leaning against the wall, standing on its side. The next day, when there was daylight, I tried to get yesterday's impression of the painting. However, I only succeeded half-ways: on its side too, I constantly recognized the objects and the fine finish of dusk was lacking. I now knew full well, that the object [objective form] harms my paintings.” [15] (p.68)

References

16. Bertamini, M. Blakemore, C. Seeing a work of art indirectly: When a reproduction is better than an indirect view, and a mirror better than a live monitor. Front Psychol, 2019 10, Article 2033. doi.org/10.3389/fpsyg.2019.02033