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

David Corwin

Abstract

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 feeling. Because eye movements are predominantly influenced by the large effects of saliency and top-down priorities, it is difficult to separate out the much smaller effect of balance. A perfectly balanced picture, a picture where balance has no effect on eye movements, might show how its absence affects movements. Given that balance is associated with a picture appearing unified and harmonious and that there is an effect known to painters and historically documented that does just that, it was thought that such pictures should be perfectly balanced. Computer models of these pictures were created by the author and were found to have bilateral quadrant luminance symmetry with a lower half lighter by a factor of ~1.07 +/- ~0.03. Based on this and the perception of a center-of-mass type effect, a weighted center of quadrant luminance calculation is proposed.

Two studies were done using the observation that a picture within a white frame can be seen as balanced when the ground is black. Comparing such an image with the same picture in a slightly different frame permits a comparison of identical images solely with respect to balance. Results show that observers, mostly painters, could distinguish picture pairs with low average pair imbalance from those high average pair imbalance. There was a significant correlation between average pair imbalance and observations that the two identical pictures appeared different. A conventional study of preference could not be done because of the necessity of using LED pictures that increase overall salience, and so decrease the aesthetic effect while retaining the effects on eye movements.

With painters who can disregard salience and with the absence of imbalance, there are no forces drawing attention to any part of the image. All parts of the picture including that in peripheral vision receive attention, and the eye seems to slide through rather than to jump from objet to object. The name given by the author to this effect is pictorial coherency. Large tonally contrasting forms with straight lines or many different forms that cannot be visually combined into a few forms prevent the effect from being seen. The percept is not seen in monochromatic red pictures where the magnocellular data stream is suppressed. That pictures can evoke such a low level percept would indicate that it belongs to much earlier evolutionary development of the visual stream where it might have been used to detect object movement and distance.

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 are said to be balanced around a geometric center or the principal pictorial axes defined by a center-of-gravity type of equation. That is said to explain both the perception of balance and the aesthetic qualities of pictures (2,3,4). Any attempt to show that low level balance is determined by some sort of equation might be associated with this erroneous concept. It is either impossible or would require too much calculation for the visual system to do this. However, in the same way that elements of salience such as color, contrast, and orientation are used to calculate the salient properties of a particular image, 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 do this 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 the 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 eye is less constrained to specific salient areas but do not quantify the effect of imbalance.

The determination of how the visual system determines pictorial balance can only be done empirically. Since there was no way to create a picture with a specific degree of imbalance, I thought that the creation of a perfectly balanced picture in which balance has no effect on eye movements could be used as a base from which other measurements might be made. In such a picture if the viewer can disregard top-down goal direction and bottom-up saliency, eye movements would be unrestrained. Because perfect balance has also been used to explain the feelings of unity and harmony evoked by some pictures, I thought that this was the explanation for a very obscure effect observed by painters that creates a remarkable feeling of unity and harmony. There is no name or metaphorical description for the percept evoked by these pictures so it is discussed only when before such a picture in art schools, particularly ones devoted to abstract painting. 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 any 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 describes 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 a percept that had no name gives his conversations a cryptic quality (12). Both Delacroix and Kandinsky give vivid descriptions of seeing such an effect in paintings (13,14).

I created a computer model of images that evoke the effect through experience acquired from many years of painting. Image analysis indicated that the percept was obtained when an LED picture, centered at eye level and parallel to the plane of vision, has bilateral quadrant luminance symmetry with the lower half being slightly lighter 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 CR2©).

A formula for balance can be obtained from this: If a 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 average quadrant luminance L replaces the mass of each object with the picture centered at (0,0):

$$X = (X_{ULQ}L_{ULQ} + X_{URQ}L_{URQ} + X_{LLQ}L_{LLQ} + X_{LRQ}L_{LRQ})/L_{TOTAL}$$
$$Y = (1.07)Y_{ULQ}L_{ULQ} + (1.07)Y_{URQ}L_{URQ} + Y_{LLQ}L_{LLQ} + Y_{LRQ}L_{LRQ})/L_{TOTAL}$$
$$L_{TOTAL} = (L_{ULQ} + L_{URQ} + L_{LLQ} + L_{LRQ})/4,$$
$$\text{Balance} = \sqrt{X^2 + Y^2} \text{ or a vector } \vec{L} = X, Y$$

Luminance L is measured on a scale from 1 to 255 with 1 being black. The 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 is very strong with a picture seen by reflected light, the percept is not as intense in LED pictures. There is a diminished feeling of unity while maintaining the effect on eye movements. Because this is more subtle, a study of subject preference could not be done. This corresponds to the observations of Bertamini et al. who noted that observers preferred a print reproduction to seeing the same picture in a mirror or on a monitor, and they preferred the mirror view of the actual image to the image on an LED screen (15). Locher et al made similar observations (16). LED images had to be used to permit precise luminance determination and reproducibility.

With respect to the effect a white framed image against a black ground is seen as one visual object. This permits the comparison of a balanced framed picture with the identical picture in a slightly different frame that renders the combination unbalanced. This also avoids the confounding variable of salience that occurs when an image is compared with even a small variant. Unfortunately, the white borders also increases the saliency of the picture. If an observer is told to compare the two images and to disregard the frame, any perceived difference would be ascribed to the effect. A study was done to show that this percept is observable. 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. A painting exhibiting coherence is 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 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 an picture is seen against a black ground.). 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 iPad image was verified to have the stated luminance balance through measurement of screen shots. The images on the iPad are approximately 5x6.75 inches and subtended visual angles of roughly 19º and 23º respectively under the conditions of the study. The observers viewed the iPad parallel to the plane of vision at arm's length and were asked whether the two central images appeared to be the same or different. 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 supplementary files).

Fig. 1 The three possible image and border combinations

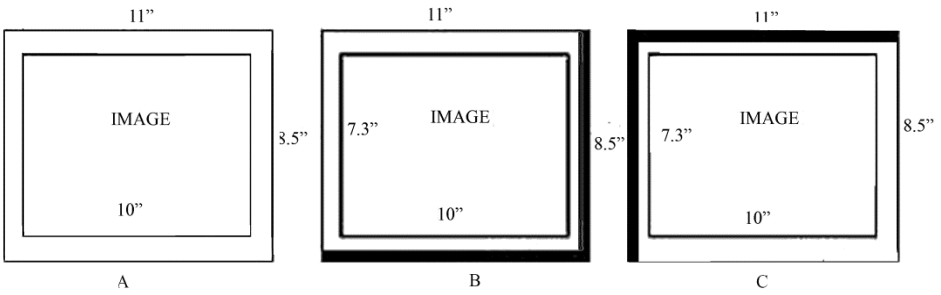


Fig. 2 Pictures used in the study pairs

The observers were predominantly artists of any age who gave their consent orally to participate in an anonymous study of how people look at pictures. No identifying information or medical history was obtained; there was no attempt to eliminate observers who were color deficient, had cataracts or impaired binocular vision. The research study was done in accordance with the Helsinki Declaration (WHO). There were no requirements other than a willingness to follow the directions to carefully view the pictures and to answer questions as to whether picture pairs appeared different or the same. 45 observers were included in the first study and 39 observers in the second study. Observers were excluded if they rapidly said they could see no difference without careful examination 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 pairs of pictures one at a time on an iPad with the examiner seated slightly behind them and asked whether the pairs appeared to be identical or different. They were told that it was thought that some pairs would be seen as the same and others as different. 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 mistake. (Table 1) Although it could not 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.

Table 1 The number of observers for each number of correct responses

1st study		2nd study	
Correct Responses	Observers	Correct Responses	Observers
10	3	10	1
9	2	9	3
8	2	8	1
7	4	7	6
6	13	6	9
5	12	5	11
4	7	4	5
3	1	3	1
2	1	2	2
Total	45		39

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 could be divided into two groups: one in which the average state of imbalance within the pair was low (the balanced pairs), and the other where the average imbalance was 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 correct identification was: $r(18) = 0.57$, $p = 0.013$ If the responses of "different" for both coherent and unbalanced pairs are compared with the average pair distance, the correlation is $r = 0.959$. (Table 2).

This might verify the hypothesis that observers see pairs as different within coherent pairs. However, it might just indicate that observers see pairs that have low average pair imbalance as different from those have 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 the eye were to see these pairs as also being different, then we could say that the eye is capable of making distinctions between low and high average pair imbalance without regard to coherency. The results indicate at a minimum that the visual system sees pictures in terms of quadrant luminance balance.

Table 2 Pair average distance and percent of observers that correctly identified the pairs (according to the author) or identified the percent of pairs seen as different.

	Coherent Pairs			Unbalanced Pairs			
	image pair	average pair distance	% correct	image pair	average pair distance	% correct	% same
1st Study	1	0.462	0.489	2	0.723	0.644	0.356
	3	0.318	0.356	4	1.894	0.711	0.289
	5	0.427	0.444	6	1.679	0.667	0.333
	7	0.509	0.378	9	0.944	0.822	0.178
	8	0.522	0.444	10	2.197	0.889	0.111
2nd Study	1	0.608	0.410	2	0.942	0.718	0.282
	3	0.652	0.359	4	2.153	0.846	0.154
	5	0.473	0.436	6	1.934	0.615	0.385
	7	0.658	0.282	9	1.389	0.744	0.256
	8	1.026	0.436	10	2.455	0.872	0.128
Mean		0.565	0.403		1.631	0.753	
StDev		0.193	0.060		0.602	0.098	

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. <001.

Pictorial balance is extremely sensitive to small differences of quadrant luminance which explains its fugacity without precise viewing conditions (table 3). Changing the lighting spectrum 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 which is normally arrived at through chance and is even then usually fleeting, and that there is no name for it, one can understand why it is not discussed in 20th century literature.

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

Pair	lateral balance	top/down	pair	lateral balance	top/down
1	1.05	0.97	**6**	**1.04**	**1.37**
	1.00	1.06		**0.97**	**1.23**
2	**1.03**	**0.95**	7	0.94	0.97
	0.99	**1.01**		1.00	1.07
3	1.00	1.06	8	1.00	1.08
	0.93	1.16		1.10	1.26
4	**0.98**	**0.90**	**9**	**1.05**	**0.87**
	1.06	**0.83**		**0.98**	**1.01**
5	0.95	1.00	**10**	**0.80**	**0.96**
	1.00	1.07		**0.86**	**1.06**

Conclusions

When an innately capable observer views a picture ignoring saliency, balance remains to direct the eye. With perfect balance there is no controlling effect on eye movements. One is attentive to the whole picture and is able to move smoothly through it instead of being attentive to one form after another. This is what Roger de Piles describes as seeing the whole picture at one time – "le tout ensemble (11). Painters view pictures differently than untrained individuals (10,17,18,19,20,). Antes et al (18) used fixation patterns to distinguish painters from non-painters, and Koide et al (21) used a prediction model of a saliency map to show that artists are less guided by local saliency than non-artists. Coherence facilitates this in an extraordinary way.

With the absence of forces drawing attention to any part of the image, all parts including that in peripheral vision receive attention.

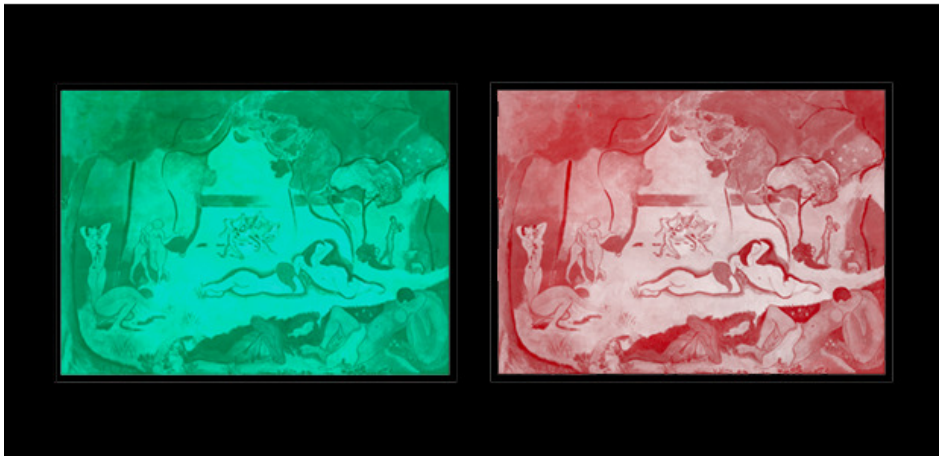
With respect to eye movements I find that it is possible for the eye to glide seemingly smoothly through a coherent painting instead of jumping from place to place. The saccades themselves seem to be easier than with an unbalanced version of the same picture. By this I mean the intangible effort it takes to lift up one's eyes from one part to another. It is simply less tiring to look at a more balanced picture, or conversely it is much more tiring to look at a very unbalanced picture. A study of how long subjects look at paintings of similar subjects differing by their degree of imbalance might show this. No one complains of visual fatigue when intently watching a football game, but they do when looking at pictures. We are accustomed to the former while unaccustomed to the latter with its pull of imbalance and so it seems to require energy.

Painters often learn to follow their eye as it moves through a picture to see where motion is impeded as a means of evaluating balance and to remove "visual obstructions." During the study an observer who was clearly familiar with the effect in his painting but who did not identify the balanced pairs particularly well, was taught to slide his eyes around. He then identified every pair correctly. It was not that he perceived the effect but was sensitive to how it affected eye movements.

The importance of peripheral vision where magnocellular ganglion cells predominate is shown by the percept disappearing in an environment of red light to which the magnocellular stream is insensitive (22,23,24). The effect is not seen in a balanced monochromatic red image but is seen in a balanced monochromatic green image of the same picture (fig. 3). Binocular vision is necessary to view the percept.

fig. 3 balanced monochromatic red and green pictures

fig. 3



There are certain types of saliency that cannot be eliminated i.e. geometric forms and many different prominent forms that cannot be combined conceptually into a few forms. The eye can learn to ignore the few major forms but cannot ignore multiple equally salient forms. 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. Coherency is best defined as the effect of the picture on eye movements.

Pictures are a very recent invention requiring the technology to create flat surfaces. They are not part of the evolutionary development of humans. Borders change an image into a picture, an object evoking the percept of balance. Since this low level percept originates in the magnocellular visual stream, the most primitive area of vision, it must have evolved for some other use. An organism using the equation of balance could represent a complex object as a simple vector and follow it extremely rapidly. The percept of balance is most likely the residual of a primitive visual system for evaluating object movement. It follows that there should be in the magnocellular stream a group of neurons that specifically evoke this percept.

Bibliography

1. Schapiro M On Some Problems in the Semiotics of Visual Art *Semiotica* 1969;1:225-242
2. Arnheim R. *Art and Visual Perception* Berkeley: University of California Press; 1974 p.10
3. Poore H R *Pictorial Composition and the Critical Judgment of Pictures*. New York:G.P Putnam Sons;1903
4. Ross, D. A *Theory of Pure Design*. Boston: Houghton, Mifflin and Co.; 1907
5. Itti L Koch C, Niebur E. A Model of Saliency-Based Visual Attention for Rapid Scene Analysis IEEE Trans Pattern Anal Mach Intell. 1998; 20(11);9 doi:10.1109/34.730558
6. Kowler E Review Eye movements: The past 25 years Vision Research. 2011;51(13): 1457-1483 doi.org/10.1016/j.visres.2010.12.014
7. Krasovskaya S, MacInnes WJ. Salience Models: A Computational Cognitive Neuroscience Review. Vision 2019;,3(4). doi: 10.3390/vision3040056.
8. Locher P The Contribution of eye movements to pictorial balance perception, a review of the literature. Empirical Studies of the Arts. 1996;14(2): 143-163
9. Locher P, Gray S, Nodine C The structural framework of pictorial balance Perception. 1996; 25: 1419-1436 doi: 10.1068/i0472aap
10. Nodine CF, Locher P, Krupinski EA The Role of formal Training on Perception and Aesthetic Judgement of Art Composition Leonardo 1993;26(3):219-227 doi: 10.2307/1575815
11. Puttfarken, T. *The Discovery of Pictorial Composition*, New Haven:Yale University; Press;2000
12. Doran, M (Ed.) Cochran JL (Trans.).*Conversations with Cézanne*. Berkeley; University of California Press ;2001 p. 88-89
13. Delacroix E, Œvres Littéraires. Études Esthétiques 4th ed Paris :G. Crés & Cie ;1923 p.67
14. Kandinsky "Rückblicke" in *Weiss P. Kandinsky in Munich*. Princeton: Princeton University Press;1979 p.68
15. 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
16. Locher P, Smith JK, Smith LF. The influence of presentation format and viewer training in the visual arts on the perception of pictorial and aesthetic qualities of paintings Perception. 2011;**30**:449-465 doi: 10.1068/p3008.
17. Long Z Peng D. Chen K Jin Z Yao L. Neural substrates in color processing: A comparison between painting majors and non-majors. Neurosci Lett, 2011;**487**: 191-195.

18. Antes J R, Kristjanson A F Discriminating Artists from Nonartists by Their Eye-Fixation Patterns. Percept Mot Skills.1991;**73**(3):893-4. doi.org/10.2466/pms.1991.73.3.893
19. Antes *JR,* Kristjanson AF The impact of level of expertise on the evaluation of original and altered versions of post-impressionistic paintings *Acta Psychologica*. 1996;**94**(2):117-131 doi.org/10.1016/0001-6918(95)00055-0
20. Zangemeister W H Sherman K Stark L. Evidence for a global scanpath strategy in viewing abstract compared with realistic images. Neuropsych 1995;**33**(8):1009-1025 doi: 10.1016/0028-3932(95)00014-t
21. Koide N Kubo T Nishida S Shibata T Ikeda K. Art Expertise Reduces Influence of Visual Salience on Fixation in Viewing Abstract-Paintings. PONE; 2015:**10**(2): e0117696. doi: 10.1371/journal.pone.0117696
22. Awasthi B, Williams MA, Friedman J. Examining the role of red background in magnocellular contribution to face perception PeerJ , 2016; **4**:e,1617 doi.org/10.7717/peerj.1617
23. Chapman C, Hoag R, Giaschi D. The effect of disrupting the human magnocellular pathway on global motion perception Vision Research, 2004; **44**(2): 2551-2557. doi.org/10.1016/j.visres.2004.06.003
24. Livingstone MS, Hubel DH Anatomy and physiology of a color system in the primate visual cortex. J Neurosci. 1984;**4**(1):309-56 doi: 10.1523/Jneurosci.04-01-00309.1984