

See discussions, stats, and author profiles for this publication at: <http://www.researchgate.net/publication/26514487>

The usefulness of eye movement recordings to subject an aesthetic episode with visual art to empirical scrutiny

ARTICLE · JANUARY 2006

Source: DOAJ

CITATIONS

18

READS

125

1 AUTHOR:



[Paul Locher](#)

Montclair State University

61 PUBLICATIONS **865** CITATIONS

SEE PROFILE

**The usefulness of eye movement recordings to subject an aesthetic episode
with visual art to empirical scrutiny**

PAUL J. LOCHER¹

Abstract

The focus of this paper is on recent findings of eye movement studies conducted by this author and his colleagues that demonstrate the valuable potential of eye-movement research to reveal the perceptual and cognitive processes that underlie an aesthetic episode with visual art.

Key words: eye-movements, aesthetics, art

¹ Correspondence concerning this article should be addressed to the author at Department of Psychology, Montclair State University, Upper Montclair, NJ 07043, or locherp@mail.montclair.edu

The usefulness of eye movement recordings to subject an aesthetic episode with visual art to empirical scrutiny

In the field of experimental aesthetics, the study of eye movements has been used as a tool by researchers for many decades to reveal covert perceptual and cognitive process that underlie the perception and aesthetic evaluation of artworks (see Nodine & Krupinski, 2003, for a review). As a viewer gathers information from an artwork such as a painting, his or her eyes move in rapid jumps or saccades followed by pauses or fixations. In general, the eyes move three to four times per second during visual search. The number, duration, and location of fixations used to scrutinize a work are a function of the interaction among the properties of a composition (e.g., its structural organization, content, style), the evaluative task being performed by the viewer (e.g., rating a composition for its complexity versus its pleasantness), and viewer characteristics (e.g., level of experience or training in the visual arts). By superimposing eye movements over a stimulus display, one can map out the location and sequence of fixations, called a scanpath, used by a viewer to encode a visual display and with this scanpath identify which pictorial features of it raised the viewer's interest and invited exploration. When used in experimental aesthetics research, a scanpath provides a graphic record of how information from an artwork was selected and processed across the time course of an individual's interaction with it, that is, during an aesthetic episode. The focus of this paper is on the way recent findings of eye movement studies conducted by this author and his colleagues provide insights into the perceptual/cognitive processes that underlie an aesthetic episode and contribute to the appreciation of paintings.

A recent advance in eye movement research is the availability of systems that permit investigators to record simultaneously a viewer's verbal reactions to a visual stimulus and his/her scanpath as the stimulus is examined. Prior to the availability of these systems, participants in eye movement studies were not permitted to move their heads during data collection because head movements interfered with calibration settings. This made it impossible for participants to verbally describe their reactions to stimuli as they looked at them. Locher (2003) has demonstrated the value of investigating the relationship between the way viewers think about an artwork, as inferred from their written reactions to it, and their ability to accurately detect original abstract and representational paintings (seen in reproduction) from two experimentally manipulated less well-organized versions of each work. Specifically, he found that hit rates for originals were highest for abstract works when participants focused on their compositional style and form and highest for representational works when their content and realism were the focus of attention. As mentioned, participants' reactions to the works were obtained from their written responses to them.

A recent study by Locher, Krupinski, Mello-Thoms, and Nodine (2005) is the first investigation to address the limitation of previous eye movement studies by recording simultaneously the eye movements and verbal reactions of observers as they evaluated for pleasingness reproductions of paintings by renowned artists. Specifically, Locher et al. examined the relationship across the time course of exploration among participants' reactions to the art stimuli obtained from their verbal responses to the works, their scanning patterns, and each painting's compositional organization. University students who had no formal training in the visual arts were shown reproductions of paintings by Bruegel, Edo Period, Giotto, Klee, Leger, Matisse, Marin, and Vermeer. Their task was to examine each painting carefully and

rate it on a 10-point scale for pleasingness. Subjects were asked to describe all of their thoughts and reactions to each picture as they looked at it and decided upon its rating.

Participants' verbal responses to each painting were classified according to seven types or levels of reactions using a classification system adapted from the one employed in the previously mentioned investigation by Locher (2003). The responses reflect a qualitative continuum of reactions ranging from attention to physical properties and pictorial elements of the compositions, to the beauty, expressiveness, style and form of compositions, and at the highest level of interaction with the art works, observers' reactions reflect personal taste as acquired by experience (either formal or informal) with art. Specifically, verbal responses were classified as reflecting reactions to: 1. single compositional elements (e.g., I see a boat); 2. several compositional elements perceived as a unit (e.g., There are men resting underneath the trees.); 3. the realism of the composition (e.g., The waves look very real.); 4. the beauty of the composition (e.g., The colors are nice and I think they are pretty.); 5. the expressiveness of the composition's content (e.g., This painting seems somewhat intimate) or the viewer's reaction (e.g., This painting makes me feel sad.); 6. the style and form of the composition (e.g., The painting has many different contrasts in color and they are not well balanced.); and, 7. statement of personal taste (e.g., I particularly like Renaissance art.). The verbal responses of each participant to each stimulus were characterized across the time course of exploration according to this reaction classification system.

Eye movements were recorded by Locher, Krupinski, Mello-Thoms, and Nodine (2005) using an ASL 4000SU Eye-Tracker with head tracking capabilities (Applied Science Labs, Bedford, MA). The art stimuli were presented individually in color on a CRT display with a pixel resolution of 1280 x 1024 and a screen size of 36 x 28 cm. This is an infrared based system that uses reflections of the pupil and cornea to sample eye-position every 1/60 of a second, generating raw eye-position data (x and y spatial coordinates and duration of fixations).

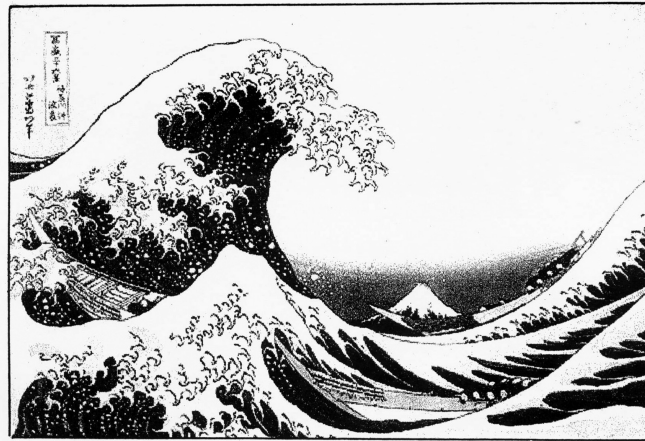
The fixation data were used to determine where and for how long (dwell time) participants focused their attention within each composition. To quantify how visual attention was allocated over each painting a 5 x 5 grid was superimposed over the display field. Each grid cell corresponded to an area of the display field that subtended a visual angle of approximately 10 degrees. By positioning a participant's scanpath for a given composition over the grid/painting image, it was possible to calculate the time he/she spent gazing at the pictorial elements in each grid location of that composition. For each participant, the distribution of dwell was calculated for each grid square in each painting. The measure of dwell time in each grid square was the cumulative fixation duration in seconds.

Analyses of the data were guided by a two-stage model of picture perception developed by this author, in collaboration with Calvin Nodine and Elizabeth Krupinski, which describes how eye movements are related to visual aesthetics (see, e.g., Locher, 1996; Locher & Nodine, 1987; Locher, Gray, & Nodine, 1996; Nodine, Locher, and Krupinski, 1993; Nodine & Krupinski, 2003). According to this model, exploration of a picture starts with a global survey of the pictorial field to acquire an initial overall impression (or gist) of the structural arrangement and semantic meaning of the composition. Based upon the gist information acquired during the first few fixations of exploration, the second phase of an aesthetic episode consists of visual scrutiny or focal analysis of presumably interesting pictorial features detected initially to satisfy cognitive curiosity and to develop aesthetic appreciation of the display.

Locher, Krupinski, Mello-Thoms, and Nodine (2005) examined participants' responses to each picture prior to 7 s of viewing time to determine how participants in their study initially reacted to the content and structural organization of the art stimuli. This value was selected because it was observed that an initial reaction could be identified within the first 7 s of 98% of all verbal protocols. Significant differences in the frequency of initial reaction types to the stimuli were found. Participants' first impressions did not reflect attention to single pictorial elements; they did not comment on the compositions' beauty or realism, nor did their reactions reflect personal taste (reaction types 1, 3, 4, and 7). Rather, their initial reactions to the artworks were to a group of pictorial elements perceived as a compositional unit, to the expressiveness of the whole composition, or to its style and form (reaction types 2, 5, and 6). The frequency of these three types of reaction across the stimulus set was 20%, 41%, and 25%, respectively. The very high occurrence of initial reactions based on information encoded from several pictorial elements within the pictorial field, especially of reactions to the expressiveness of the pictorial content, supports the view that perception of the art stimuli began with a response to the holistic organization of the works.

Visual coverage of the pictorial fields of the eight art stimuli for the first 7 s of exploration was very similar across paintings; participants directed their gaze to 32% (i.e., 8 of 25 grid cells), on average, of the pictorial fields. Figure 1 contains a picture of one of the art stimuli used by Locher, Krupinski, Mello-Thoms, and Nodine (2005), viz., *The Great Wave at Kanagawa* from the Edo Period. Below the composition is the distribution of gaze densities in seconds for each grid location. The distributions are shown for the first 7 s of the aesthetic episode and for the total time of exploration, the upper and lower values in each cell, respectively. Cells with no values are areas of the pictorial field that received either no fixations or received less than 3 standard errors of dwell time during the aesthetic episode. As seen in the figure, only 7 of the 25 (28%) grid locations were fixated during the first 7 s of exploration. Furthermore, viewers' attention was drawn initially to regions of the composition that contain major compositional elements. Thus, the initial holistic responses, which accounted for 81% of all initial responses to this picture, were generated from direct fixation of compositional elements contained in approximately one-third of the pictorial field. Furthermore, participants typically began to describe their reactions to the pictures after approximately 2 s of exploration, or approximately 6 fixations. These findings provide support for the view that an aesthetic episode begins with the generation of a holistic or global impression of the structure of a composition.

As mentioned, according to the two-stage model of picture perception described above, the second phase of interaction with a picture consists of visual scrutiny of pictorial features which are fit into a meaningful visual concept as a viewer develops an aesthetic appreciation of the picture. Analyses of the nature of participants' responses made after the first 7 s of viewing to the end of the time course of exploration revealed that the frequencies of reaction types were relatively similar during the two stages of exploration. Specifically, participants' verbal responses to the artworks during the second stage indicated that they reacted to groups of elements perceived as a compositional unit, to the expressiveness of the whole composition, or to its style and forms (reaction types 2, 5, and 6). One notable difference between participants' reactions during the two stages of exploration was that the distribution of these three types of reactions during the second phase of processing (viz., 26%, 30%, 25%) was not reliably different, whereas participants reacted significantly more to the pictures' expressiveness during the initial stage of exploration (viz., 41% vs. 26%, respectively).



		--	--	
		3.9	3.1	
4.8	8.7	6.4	--	
4.9	4.6	6.5	11.1	
15.5	17.2	--		
5.4	11.8	6.5		
27.4	7.1	--		
3.7	16.9	8.9		
--	--			
2.9	3.3			

Figure 1:

Percent viewing time for grid locations of the Edo Period painting *The Great Wave at Kanagawa* for the first 7 seconds of viewing (upper value) and total viewing time (lower value).

Percent visual coverage of the compositions' pictorial fields for total viewing time averaged 61% across the stimulus set. As reported above, average coverage during the initial stage of exploration was 32%. Thus, only an additional 29% of the compositions' areas received direct attention as measured by gaze dwells following participants' initial reactions to the art stimuli. Shown in Figure 1 is the percent of total viewing time in each of the 25 grid cells for the Edo Period painting *The Great Wave at Kanagawa*. Note that viewing times for the entire aesthetic episode are shown in only 14 of the 25 cells. This indicates that the gaze was directly focused on only 56% of this composition's pictorial field during the total time participants spent looking at it. Note also that participants concentrated their gaze in the left pictorial field of this composition which contains the elements that are the principal contributors to the perceptual and semantic meaning of the work. Scanpaths for the other

pictures included in the stimulus set reflect a similar concentration of the gaze on the pictorial elements that contribute to the organization and interest of the works. Clearly not all of the pictorial elements of a painting contribute to the same extent to the structural organization that conveys a composition's main theme. Ultimately, an understanding of the motivational effects of a painting must include an explanation of the way the assemblage of its pictorial elements and their relationship to the unified whole transmit "aesthetic information" to a viewer. Computer graphics systems now make it possible to manipulate the structural organization of a painting, such as its balance, so that the original and altered versions can not be distinguished one from the other except in ways manipulated. Needed, then, to obtain a more comprehensive understanding of the nature of the perceptual/cognitive processes that underlie an aesthetic episode are eye movement studies that simultaneously record viewers' fixation patterns and verbal reactions to original compositions and to their structurally altered perturbations. This is the future direction of research in our laboratory.

One structural characteristic of artworks that *has* received considerable attention in the field of experimental aesthetics is balance. We turn now to recent research by this author and his colleagues that demonstrates the usefulness of eye-movement studies in providing empirical support for theoretical writings concerning the influence of pictorial balance on perceptual processing of pictures.

For centuries artists and writers on Western art have asserted that balance is the primary design principle for unifying structural elements of a composition into a cohesive perceptual and narrative whole that creates the essential integrity or meaning of the work. One feature of a composition that is a major contributing factor to its balance structure is the distribution of "weight" in the composition. The perceived weight of a pictorial element within the structural organization is determined by its size, shape, location within a composition, and its implied "directionality." Additionally, artists have known for a long time that the juxtaposition of colors within a composition is a major contributing factor to its balance or harmony. A general principal concerning the use of color found in treatises on composition is that a large area of a dull, unsaturated color can balance a small area of a highly saturated color. This color-weight principle of color balance was quantified by Munsell (1905) in his law of inverse ratios of areas which states that the areas of colors used in combination should be inversely proportional to the product of their values (brightness) and chromas (saturation).

There is a long history of experimental investigations that support theoretical writings concerning the perceived color balance between pairs of color. The interactive combination of colors typically found in everyday color displays, such as paintings, pictures and computer images, had not been investigated prior to a recent study by Locher, Overbeeke, and Stappers (2005). They examined the interactive contribution of the color and size of the three areas occupied by the primary colors red, blue, and yellow in adaptations of abstract compositions by Mondrian to the perceived weight of the areas and the location of the balance centers of the compositions. The paintings that served as stimuli for this research each consisted of vertical and horizontal black lines extending either to the edge of the pictorial field or to another black line, and three solid color areas on a white background. This compositional organization can be seen in the rendition of Mondrian's painting entitled *Composition with red, yellow, and blue* (1929) shown in Figure 2. Locher et al. created their art stimuli by experimentally changing the colors in the three areas of six original works by the artist so that the resulting five variations and the original constituted the six possible spatial arrangements of the three colors in the three locations. Participants determined the location of bal-

ance center of each of the 36 stimuli and rated the apparent weight or heaviness of each color area. It was found that for both art-trained and untrained participants the perceived weight of a color, especially red and yellow, varied as function of the size of the area it occupied. Specifically, the largest color area in each composition was rated significantly heavier when red than yellow. When blue, the largest area's average weight fell between the weights as-

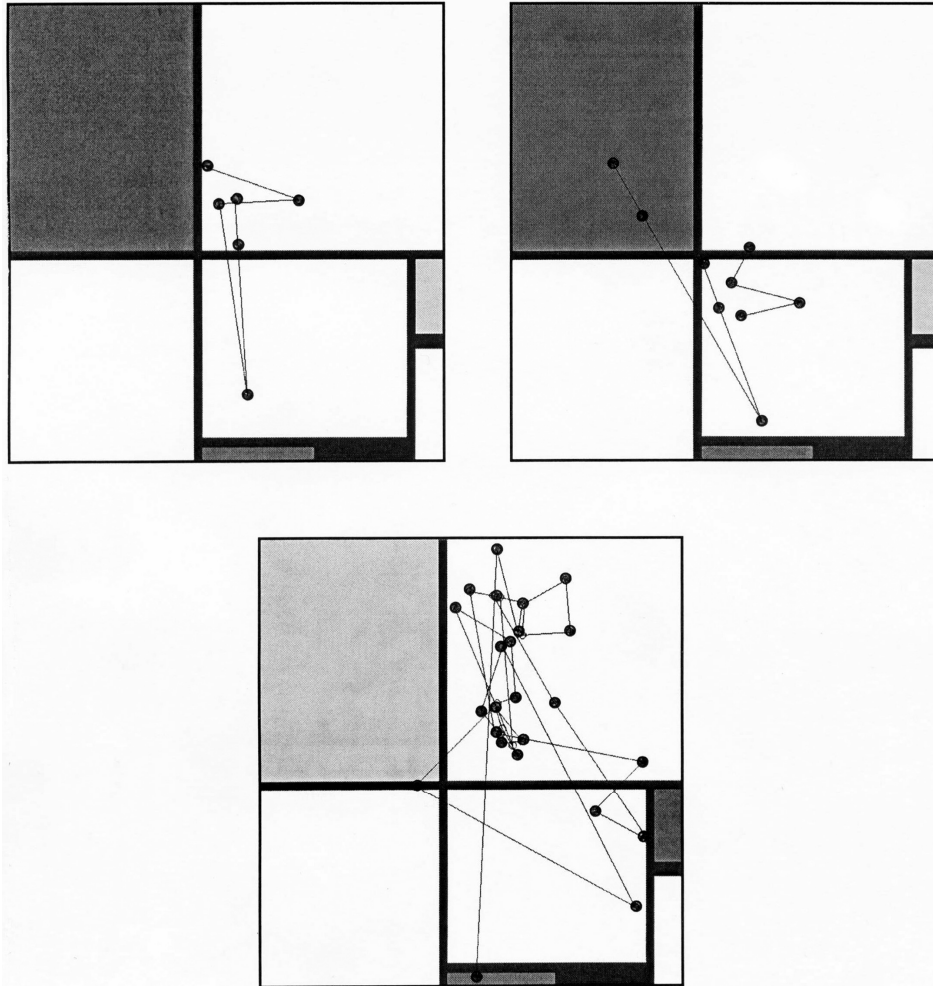


Figure 2:

Renditions of Mondrian's painting entitled *Composition with red, yellow and blue* (1929) with the scanpath of a different viewer superimposed on each version. Each scanpath consists of fixation positions (dots) connected by lines indicating the sequence of fixations. The color of the large shaded area in the upper left region of the pictorial field was red, blue, or yellow in the top left, top right, and bottom compositions shown, respectively.

signed when the area was either red or yellow. Average differences in weight between the two smaller color areas were not reliable. Additionally, participants perceived significant shifts in the locations of the balance centers between the originals and their altered versions.

Locher, Overbeeke, and Stappers' (2005) findings do not provide direct evidence that the observed differences in the perceived weights of the colored areas and the location of the balance centers in the Mondrian stimuli are the result of differences in viewers' scanpaths as they performed the task. Locher, Dolese, and Guise (2005) sought such evidence. Specifically, they asked participants who had no training in the visual arts to determine the balance center of three versions of each Mondrian painting studied by Locher, Overbeeke, and Stappers in which the largest color area was either red, yellow, or blue. Participants saw only one version of each Mondrian composition, and these were embedded in a stimulus set that included other abstract compositions. Their task was to determine the location of the balance center of each composition as their eye movements were recorded.

Eye-movement data for the original and modified versions of the paintings were examined to determine if altering the color-balance of the original compositions affected the way participants viewed the paintings. To do this, the scanpath for each observer for each stimulus was quantified by superimposing a 5 x 5 grid over it and measuring the percentage of total fixation time (gaze density) in each grid location, as explained above for the study by Locher, Krupinski, Mello-Thoms, and Nodine (2005). To determine whether a significant difference in the distribution of the gaze existed between the three versions of each composition, percentage fixation data were tested for significance with a version by grid cell analysis. It was found that observers explored the three versions of a given composition differently when its largest color area was red versus blue versus yellow. The nature of these differences reflected not only changes in color but also the overall pictorial organization of each art stimulus. An example of the influence of the color manipulation on visual exploration of the art stimuli can be seen in Figure 2. Shown in the figure are the scanpaths of different participants superimposed on each of the three color variants of Mondrian's *Composition with red, yellow, and blue* (1929). Differences in the distribution of participants' gaze over the pictorial field of each version are clear. Note that the participant who viewed the composition when the largest color area was red concentrated his fixations slightly above the center of the pictorial field whereas when that area was blue, a second viewer concentrated her gaze just below the physical center. A third viewer used a very different scanpath to explore the composition when the largest area was yellow. As seen in Figure 2, more fixations were used to explore the pictorial field as this person decided upon the location of the balance center and these were distributed to a large area of the work. Two regions of concentrated dwell time can be seen; one in the center of the composition and another above this area.

Locher, Overbeeke, and Stappers (2005) observed that the perceptual weight for the largest area of this composition was perceived as significantly heavier on average when red than when yellow, 7.95 versus 7.18, respectively, on a 10-point scale. When this area was blue, viewers rated its perceptual weight as 7.76, on average, a value not significantly different than the other two means. Differences in the scanpaths for the composition shown in Figure 2 demonstrate that the perceived weight of compositional elements of abstract art are related to the way viewers explore a composition. Quantitative analyses conducted on the full set of stimuli revealed the influence of color on participants' scanpaths and their perception of the location of a composition's balance center thereby providing empirical evidence that altering

the color-balance of the original abstract compositions affected the way viewers visually explored them.

In sum, the research findings described in this paper demonstrate the valuable potential of eye-movement research to reveal attention and cognitive processes that occur during an aesthetic episode and that contribute to the aesthetic evaluative judgments of art works. Together, state-of-the-art eye-fixation recording systems and computer graphics programs make possible a level of experimental sophistication heretofore unavailable to the scientific study of the visual arts.

References

1. Locher, P. J. (1996). The contribution of eye-movement research to an understanding of the nature of pictorial balance perception: A review of the literature. *Empirical Studies of the Arts*, 14, 143-163.
2. Locher, P. (2003). An empirical investigation of the visual rightness theory of picture perception. *Acta Psychologica*, 114, 147-164.
3. Locher, P., Dolese, M., & Guise, K. (2005). The perceived weight of color: Its effects on the perception of pictorial balance. Manuscript in preparation.
4. Locher, P., Gray, S., & Nodine, C. (1996). The structural framework of pictorial balance. *Perception*, 25, 1419-1436.
5. Locher, P., Krupinski, E., Mello-Thoms, C., & Nodine, C. (2005). The time course of perception during an aesthetic episode: Manuscript in preparation.
6. Locher, P., & Nodine, C. (1987). Symmetry catches the eye. In J. K. O'Regan & A. Levy-Schoen (Eds.), *Eye movements: From physiology to cognition* (pp. 353-361). Amsterdam: Elsevier Science Publishers.
7. Locher, P., Overbeeke, K., & Stappers, P. J. (2005). Spatial balance of color triads in the abstract art of Piet Mondrian. *Perception*, 34, 169-189.
8. Munsell, A. H. (1905). *A color notation*. Boston, MA: George H. Ellis Co.
9. Nodine, C., & Krupinski, E. (2003). How do viewers look at artworks? *Bulletin of Psychology and the Arts*, 4, 65-68.
10. Nodine, C., Locher, P., & Krupinski, E. (1993). The role of formal art training on the perception and aesthetic judgment of art compositions. *Leonardo*, 26, 219-227.