Spatial Vision, Vol. 21, No. 1–2, pp. 55–77 (2007)

Also available online - www.brill.nl/sv

Visual interest in pictorial art during an aesthetic experience

PAUL LOCHER [1,*], ELIZABETH A. KRUPINSKI [2], CLAUDIA MELLO-THOMS [3] and CALVIN F. NODINE [4]

[1] *Department of Psychology, Montclair State University, Montclair, NJ 07043, USA*
[2] *Department of Psychology, Telemedicine and Radiology, University of Arizona, Tucson, AZ 85721, USA*
[3] *Department of Radiology and Training Program of Biomedical Informatics, University of Pittsburgh, Pittsburgh, PA 15213, USA*
[4] *Department of Radiology, University of Pennsylvania, Philadelphia, PA 19104, USA*

Received 31 May 2006; accepted 26 December 2006

Abstract—Two experiments were performed during which adults untrained in the visual arts were shown digital versions of eight paintings by renowned artists. In Experiment 1 participants' written reactions following a single 100 ms glance at each work were found to overwhelmingly reflect an initial holistic impression (i.e. gist) of the structural arrangement and semantic meaning of the paintings. In the second experiment participants' eye movements and verbal reactions were recorded as they evaluated each reproduction for pleasingness. Analyses reveal the relationships between the content and structural organization of the art stimuli and the way viewers select, process and think about information contained in paintings across the time course of an aesthetic experience. The results are interpreted in terms of an information-processing stage model of visual aesthetics according to which perceptual-cognitive processing of an art stimulus begins with the rapid generation of a gist reaction followed by scrutiny of pictorial features directed in a top-down fashion by cognitively-based evaluative processes.

Keywords: Pictorial art; gist reaction; visual exploration; verbal reactions; aesthetic evaluation.

INTRODUCTION

Visitors looking at paintings in museum galleries glance at a composition and either almost immediately move on to another work or stop and spend time with it (Smith and Smith, 2001, 2003). Museum visitor behavior is in accord with our two-stage model that describes the relationship between eye movements and visual aesthetics (Locher, 1996; Locher and Nodine, 1987; Locher *et al*., 1996; Nodine

*To whom correspondence should be addressed. E-mail: locherp@mail.montclair.edu

and Krupinski, 2003; Nodine *et al*., 1993). According to the model, exploration of a painting starts with a rapid global survey of the pictorial field to acquire an initial overall impression (or gist) of the structural arrangement and semantic meaning of the composition. The gist information acquired during the first few fixations of exploration drives the second phase of an aesthetic experience that consists of visual scrutiny of interesting pictorial features detected initially to satisfy cognitive curiosity and to develop aesthetic appreciation of the display. The general purpose of the present research was to subject this two-stage processing model to further empirical scrutiny and expand knowledge concerning the relationship between the pictorial content and structural organization of paintings and the way individuals visually select and think about this information across the time course of the aesthetic experience.

Two experiments were conducted. The first one was designed to investigate the types of cognitive content that constitute a gist reaction to art. It was hypothesized that the majority of participants' gist reactions to tachistoscopically presented reproductions of paintings by renowned artists would reflect analyses of a composition's artistic, emotional, or aesthetic qualities, factors that might cause one to want to spend more time with the work, rather than attention to its individual pictorial elements or local features. This finding would indicate that participants' initial reactions were based on Rasche and Koch's (2002) explanation of the nature of a gist response and the neural mechanisms responsible for it. Briefly, they argue that a gist response to an image is generated from visual input that is concurrently spread across many cortical areas which then communicate with each other rapidly in a distributive (rather than hierarchical) manner, quite likely with the help of interpretation (feedback), and possibly even before the first saccade (Kundel *et al*., 2007).

The second experiment tested the notion that gist recognition is based only on a subset of image information by recording eye fixations as individuals looked at artworks used in Experiment 1 and evaluated them for pleasingness. The study of eye movements has proven to be a very useful tool to investigate how the pictorial composition of artworks affects viewing behavior during an aesthetic experience (see Locher, 1996; Nodine and Krupinski, 2003, for reviews of this literature; as well as Locher and Nodine, 1987; Nodine and McGinnis, 1983). In addition to recording participants' eye movements, a concurrent verbalization procedure was used in the second experiment to get subjects to verbalize their thoughts as they formulated their evaluative judgment of the work's pleasingness. The advantage of knowing the way individuals visually explore *and* think about a picture in the present study provides greater insights into the perceptual-cognitive processes underlying an aesthetic encounter with art stimuli than does information concerning just one of these behaviors (see Ericsson and Simon, 1993, for a review of this literature).

It is apparent that when one looks at a painting not all of its pictorial elements contribute to the same extent to the structural organization that conveys the

composition's main theme. Artists typically concentrate elements of major interest in the center of the pictorial field, especially in representational works (Arnheim, 1988; Bouleau, 1980). An optically based explanation has frequently been put forth to explain this aspect of compositional structure (see for example Bearden and Holty, 1969). This view asserts that much of an artist's search for the proper scale for paintings is conditioned by the fact that visual acuity is best for pictorial features focused in the center of the visual field and drops off substantially away from its center. It is to be expected, therefore, that some features of a picture will receive a high degree of attention and other regions will go unattended during both the initial and second stage of processing. Eye-movement studies have, in fact, consistently shown such differential exploration of paintings and pictures (Buswell, 1935; Locher *et al*., 1993, 1996; Molnar, 1981). For example, Locher *et al*. (1996) observed that adults whose eye movements were recorded as they evaluated representational and abstract paintings for balance paid very little or no attention to the outer regions of the paintings. They concentrated their gaze (approximately 65% of total fixation time) in the middle 25% of the pictorial field. We anticipated that observers in the present study would direct their gaze to pictorial features or areas of the artworks that were important components of the compositions' structural skeletons, as identified by experts (see below), while other regions of lesser structural and semantic importance would go unnoticed.

In sum, the present research subjected a two-stage processing model of visual aesthetics to further empirical scrutiny by studying the relationship between the pictorial content of paintings and the way individuals visually grasp, explore and think about this information across the time course of an aesthetic experience.

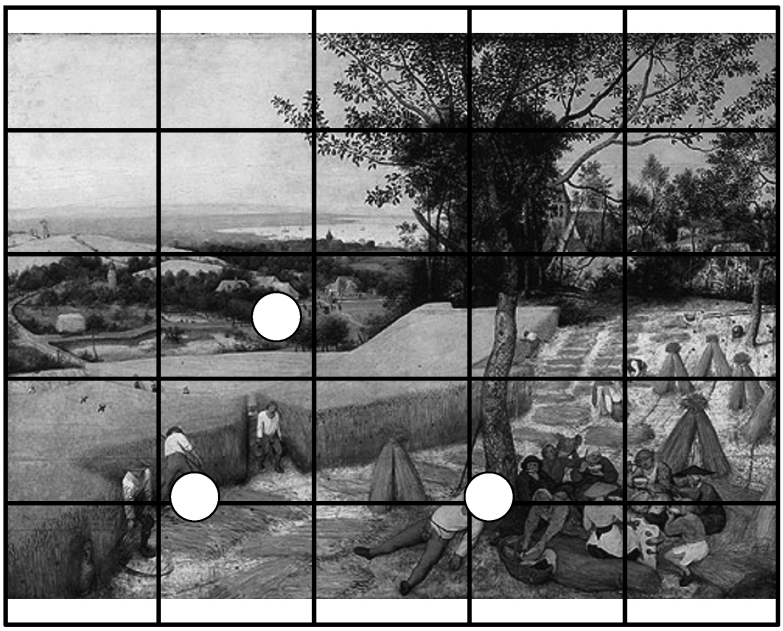
EXPERIMENT 1

Method

Participants. Twenty female and nine male graduate students who reported no formal education or studio training in the visual arts volunteered to participate in this study. They ranged in age from 21 to 42 years ($M = 26.8$, $SD = 4.7$).

Stimuli. The art stimuli used in Experiments 1 and 2 consisted of digital versions of eight paintings by renowned artists downloaded from The Metropolitan Museum of Art, New York City collection web site (http://www.metmuseum.org/ accessed January 2004). They comprised Bruegel's *The Harvesters*, Edo Period *The Great Wave at Kanagawa*, Giotto's *The Epiphany*, Klee's *Temple Gardens*, Leger's *Woman with a Cat*, Marin's *Brooklyn Bridge*, Matisse's *Promenade among the Olive Trees* and Vermeer's *Young Woman with a Water Pitcher*. The set of artworks, shown in Fig. 1a–h, respectively, were selected to represent a range of styles along the abstract-representational continuum.

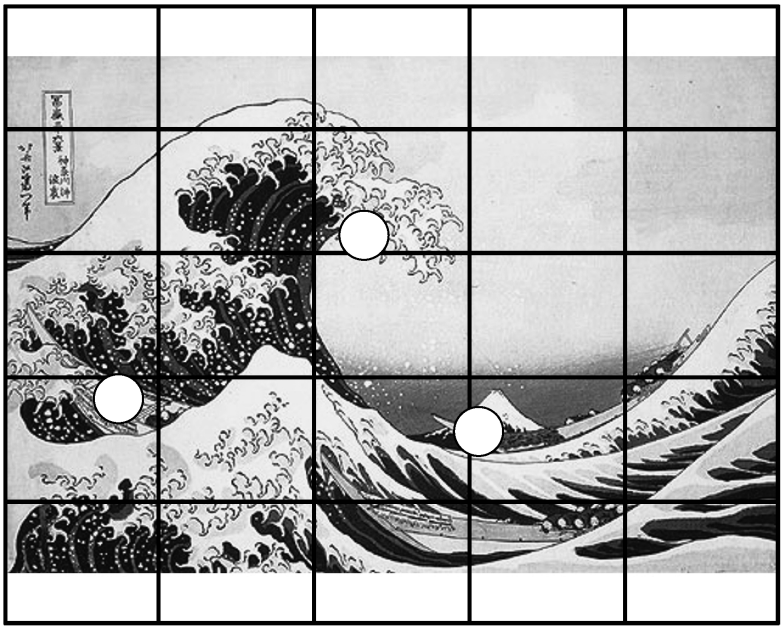
In Experiment 1 the art stimuli were projected within a 75 in. (178 cm) square viewing area in the middle of a screen using a Kodak slide projector fitted with



		9,7,9	11,14,9	0,9,0
	9,10,0		11,0,8	
0,0,7	0,5,9	17,7,6	0,0,6	
0,9,0	0,8,7	0,0,7	0,0,8	

(a)

Figure 1. Percent viewing times (dwell times) for grid locations for the first 3 s of viewing (first value), period 3–7 s (second value), and 7 s to total viewing time (third value) for each artwork. The artworks are: (a) Bruegel's *The Harvesters*, (b) Edo Period *The Great Wave at Kanagawa*, (c) Giotto's *The Epiphany*, (d) Klee's *Temple Gardens*, (e) Leger's *Woman with a Cat*, (f) Marin's *Brooklyn Bridge*, (g) Matisse's *Promenade among the Olive Trees*, and (h) Vermeer's *Young Woman with a Water Pitcher*. The three major aspects of the structural organization of each artwork are shown as white dots on its image.

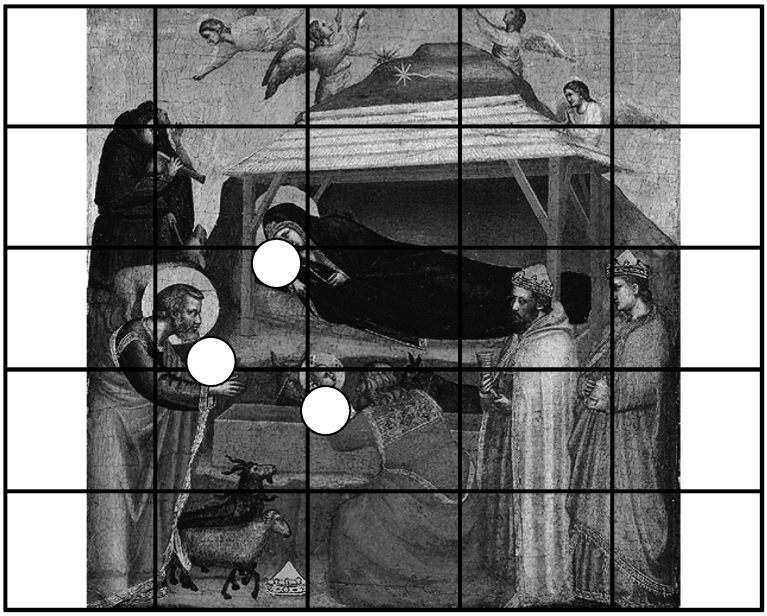


0,7,5	**5,0,0**	**8,8,6**	**9,6,12**	
0,0,6	**20,7,11**	**15,8,0**		
	20,28,14	**0,8,9**		
6,0,0	**0,0,3**			

(b)

Figure 1. (Continued.)

a Lafayette Instrument Company tachistoscopic lens (model 43016) controlled by a Lafayette Instrument Company timer (model 43-11-15). Each image was projected to fill as much of the viewing area as possible, leaving the surround white. The largest painting in actual size, Bruegel's *The Harvesters* (Fig. 1a), measured 125 cm × 170 cm on the screen, the projected image of the smallest painting in actual size, Klee's *Temple Gardens* (Fig. 1d), measured 116 cm × 168 cm. Groups



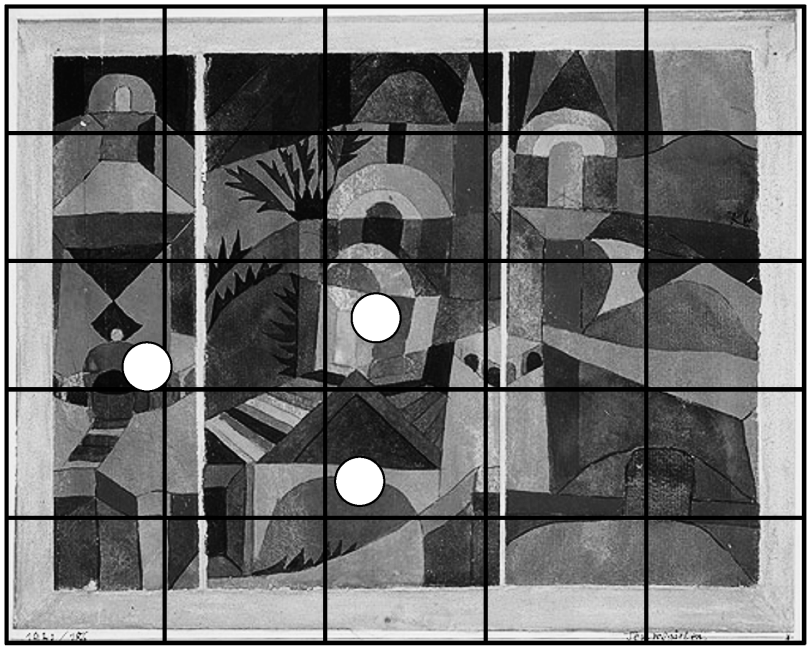
	0,0,8	**8,9,12**		
	13,15,11	**22,13,10**	**0,7,8**	
	9,11,9	**8,7,5**		
	0,0,7	**8,14,6**		

(c)

Figure 1. (Continued.)

of four or five participants sat at a distance of approximately 10 feet (274 cm) from the screen. At this distance, the viewing area of the screen subtended approximately 34° visual angle horizontally and vertically.

Procedure. The study received IRB approval from the Montclair State University Human Subjects Committee. Each session lasted approximately 30 min. The experiment was performed in a small classroom in which the lighting was very dim,

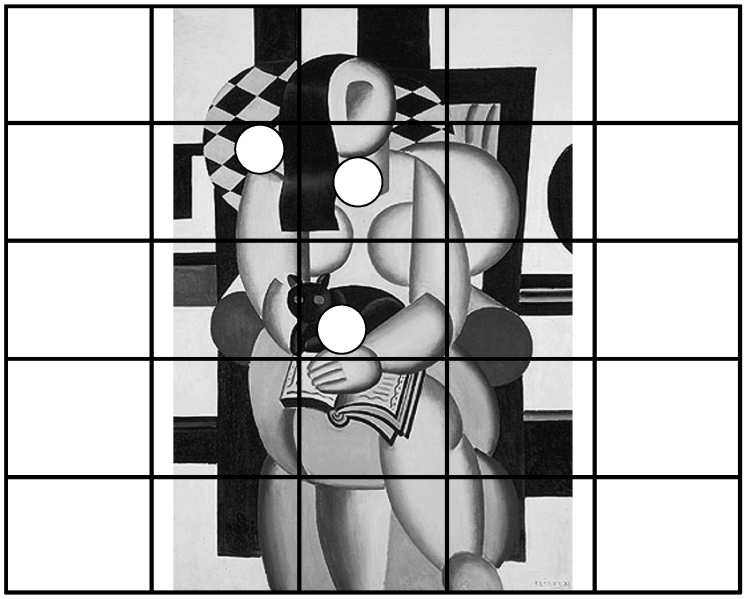


0,5,0		0,0,5		
9,9,9	9,0,7	20,20,12	6,9,8	
	10,16,8	18,12,11	0,0,5	

(d)

Figure 1. (Continued.)

but sufficient for participants to write their reactions to the stimuli. Each group of participants saw a different random order of the eight art stimuli. Following each 100 ms presentation they wrote five descriptions and/or impressions of the reproductions they would tell someone who had not seen the artwork in an attempt to describe the work to that person. Each participant was given a 10-page booklet in



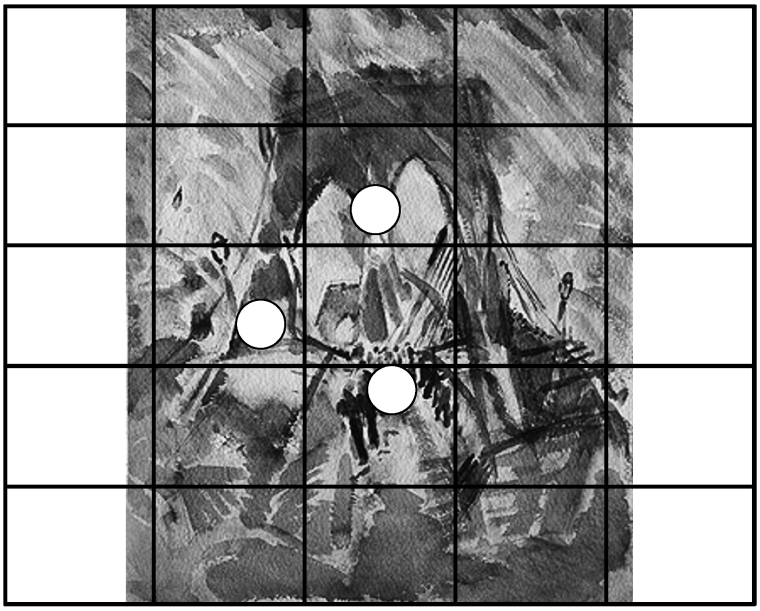
	0,5,0	**13,10,10**		
	14,0,9	**24,15,18**		
	0,5,5	**16,24,15**		
	0,10,12	**17,16,17**		

(e)

Figure 1. (Continued.)

which to record the following information. Demographic information (gender, age, training of any sort in the visual arts) was recorded on the first page. Pages 2 through 9 each contained the identification number of the stimulus seen in a given trial and the numbers 1 through 5 spaced along the left margin of the page for participants to write their reactions to the artwork.

After the participants' reactions to the eight artworks had been collected, the stimuli were presented individually a second time for 100 ms in the same order



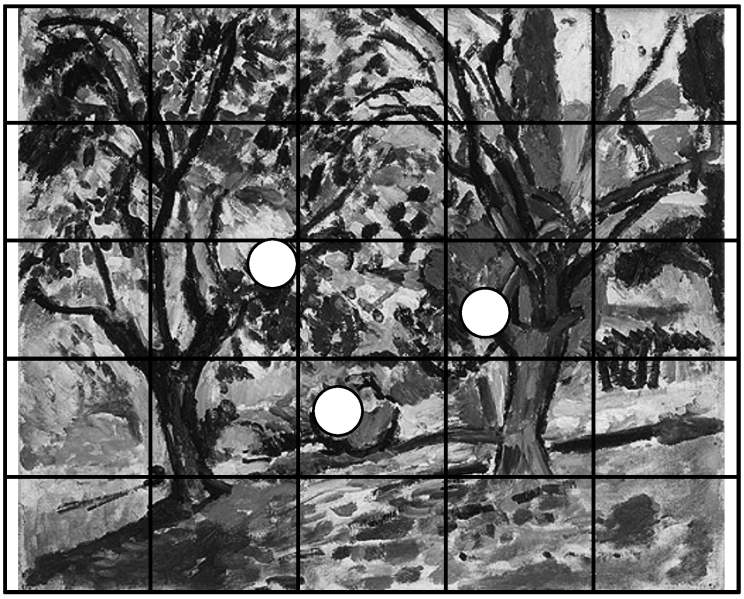
		13,15,12	0,5,0	
	0,0,5	48,26,7	7,0,10	
	0,5,9	13,20,7	0,0,5	
		0,9,0		

(f)

Figure 1. (Continued.)

and participants rated the pleasingness of each one on a 10-point scale (1 = not at all pleasing and 10 = very pleasing) in the appropriate space on the tenth page of the booklet.

Classification of verbal reactions to the art stimuli. Participants' five written statements about each painting were classified according to a system developed by Locher (2003) and his colleagues (Lega *et al.*, 2003) to characterize adult viewers'

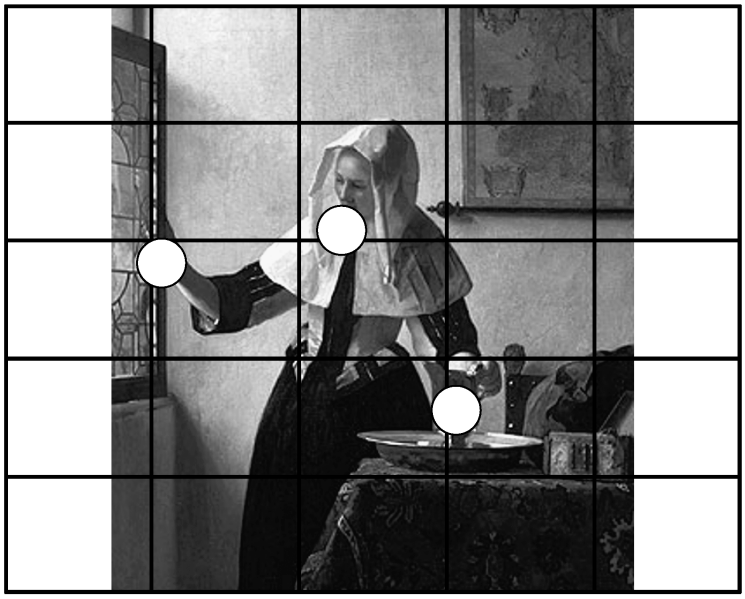


0,6,0		12,13,12	0,8,0	
	6,7,5	18,9,7	11,10,10	5,0,0
	7,10,9	0,6,7	0,0,5	
	7,0,0			

(g)

Figure 1. (Continued.)

reactions to paintings. The classification system consists of six types of reactions that reflect a qualitative continuum of responses ranging from attention to physical properties and pictorial elements of the compositions, to the beauty, expressiveness, style and form of the compositions. Specifically, each written statement was classified as reflecting reactions to:



		12,15,7	**6,12,10**	
0,6,0	**0,0,5**	**19,12,12**		
0,5,0	**16,15,11**	**30,22,22**		

(h)

Figure 1. (Continued.)

1. Single compositional elements (e.g. I see a boat. You can see the outline of a chair. The first thing I noticed was the cat. It looks like a bridge.) (Examples given for each type of reaction are participant comments.)
2. Several compositional elements perceived as a unit (e.g. There is a woman reading a book. There are men resting underneath the tree. There are people walking in the background. There is a large wave with breaking foam.)

3. The realism of the composition (e.g. The forest looks very real. I am not sure what this is supposed to be. The painting seems very natural.)
4. The beauty of the composition (e.g. It's very pretty. It's not awful. It's nice. The colors are warm and I think they are pretty. The browns and blues are a pleasing mix of colors.)
5. The expressiveness of the composition's content (e.g. It says something about the forces of nature. This painting seems somewhat intimate. It's rather depressing. The angels above show God's creation. The sky is mysterious.) or of the viewer's reaction (e.g. The colors are dreary and dull and make me feel sad. It's not very exciting to me. This one is really cool. There is an intense dark feeling conveyed.)
6. The style and form of the composition (e.g. The painting is very abstract. The entire painting is composed of squares and triangles. This painting has many different contrasts in color. I think it is over-styled and flat. It has geometrical shapes formed together into three panels. Well, this one is just all over the place.)

Results and discussion

First reactions to the art stimuli. The first author and his research assistant separately classified each participant's five reactions to each artwork according to the system described above. Inter-rater agreement was 93%. Discrepancies were discussed by the two raters and a single classification was agreed upon for each reaction. We examined the distribution of reaction types using a separate chi-squared analysis for the initial response to each picture and for the second, third, fourth and fifth responses combined. As shown in the Table 1, the percentages for both distributions were significantly different. Specifically, few of participants' first reactions to the compositions or their additional comments (2% and 4% respectively) reflect attention to single compositional elements. Rather the majority

Table 1.
Percentage of each reaction type for all eight artworks elicited during Experiments 1 and 2

	Reaction type							
	1	2	3	4	5	6	$X^2(6)$	p
Experiment 1								
Initial comment	2	25	14	4	21	34	45.41	0.01
Additional 4 comments	4	28	6	7	28	27	43.03	0.01
Experiment 2								
Responses prior to 7 s	6	21	0	6	41	26	70.80	0.01
Responses after 7 s	5	29	3	5	27	31	54.33	0.01

Note: Reaction type: 1 = single compositional elements, 2 = several compositional elements perceived as a unit, 3 = realism of the composition, 4 = beauty of the composition, 5 = expressiveness, 6 = style and form of the composition. A separate Chi-square was performed on each row of percentages and the obtained statistics are reported in the table.

Table 2.
Mean percent coverage of the eight pictorial fields for three viewing periods; mean viewing time in seconds, and mean pleasingness ratings for Experiments 1 and 2 for the eight paintings

Artist	Percent coverage for three viewing periods			Viewing time	Pleasingness ratings	
	First 3 s	First 7 s	Full viewing		Exp. 1	Exp. 2
Bruegel	0.20 (0.07)	0.36 (0.03)	0.52 (0.04)	36.1 (3.3)	5.92 (0.61)	4.46 (0.28)
Edo Period	0.28 (0.02)	0.36 (0.04)	0.44 (0.05)	33.7 (2.7)	4.72 (0.81)	6.42 (0.37)
Giotto	0.24 (0.03)	0.28 (0.04)	0.36 (0.07)	33.3 (2.2)	3.37 (0.87)	6.13 (0.54)
Klee	0.24 (0.04)	0.28 (0.05)	0.36 (0.05)	32.6 (2.6)	4.17 (0.99)	6.60 (0.47)
Leger	0.33 (0.04)	0.53 (0.03)	0.53 (0.04)	30.8 (3.6)	3.41 (0.79)	7.13 (0.52)
Marin	0.27 (0.03)	0.47 (0.04)	0.60 (0.03)	31.0 (2.6)	4.27 (0.82)	6.33 (0.56)
Matisse	0.28 (0.05)	0.36 (0.05)	0.44 (0.04)	33.0 (3.1)	6.00 (0.78)	5.66 (0.37)
Vermeer	0.33 (0.03)	0.40 (0.03)	0.40 (0.03)	29.9 (2.2)	4.89 (0.86)	6.64 (0.60)
M_s	0.26	0.38	0.46	32.5		

Note: Values in parentheses are standard errors. Pleasingness ratings range from 1 = not at all pleasing to 10 = very pleasing.

of participants' initial and later reactions to the artworks reflect attention to a group of pictorial elements perceived as a compositional unit, to the expressiveness of the whole composition, or to its style and form (i.e. reaction types 2, 5 and 6, respectively). This indicates that observers had early access to pictorial information distributed across the pictorial field and is consistent with our two-stage processing model (Nodine and Krupinski, 2003) and Rasche and Koch's (2002) view of the gist response.

Pleasingness ratings. The average pleasingness ratings for eight artworks obtained from participants following a brief glance at each one are presented in Table 2. These data will be discussed later in this paper in conjunction with ratings obtained in Experiment 2 from participants who were permitted unlimited viewing of the works.

EXPERIMENT 2

Method

Participants. Eight female and seven male undergraduate students who reported no formal education or studio training in the visual arts volunteered to participate in this study. They ranged in age from 18 to 27 years ($M = 21.7$, $SD = 2.7$). Subjects received extra credit points in a Psychology class for participating. All had 20/20 or corrected vision. None had a lazy eye or any other visual abnormality that would interfere with the eye-tracking procedure. None reported color vision disorders.

Stimuli. A PowerPoint presentation (Microsoft 2000) was created with the eight art images, with each image centered for viewing as a slideshow on a color computer

monitor (Hitachi CM771, 1280 × 1024 pixels, 32-bit, flat screen). The viewing area was 36.5 cm × 27.5 cm. The size of the image on the screen of the compositions by Bruegel, Edo Period, Klee, and Matisse was 36 cm × 25 cm, it was 25.5 cm × 26.0 cm for the work by Giotto, 17.5 cm × 25.0 cm for that by Leger, 22.0 cm × 25.0 cm for the Marin work, and 23.0 cm × 25.0 cm for the one by Vermeer.

Procedure. The study received IRB approval from the University of Arizona Human Subjects Committee. Each session lasted approximately 45 min. The participants were told they would be shown a series of eight paintings on the computer monitor, each one for unlimited viewing time, while their eye-position was recorded. As they viewed each painting they were to talk out loud about their reactions to and thoughts about the painting. In evaluating a painting, participants were instructed to try to adopt both an intellectual viewpoint emphasizing objective analysis of the pictorial elements and the compositional arrangement within the painting and a subjective, personally based analysis of their own likes and dislikes of the content. In responding, they were asked to make their verbal descriptions of their perceptions and evaluations of each painting as detailed as possible.

The verbal reports were recorded on a standard tape recorder. When participants finished verbalizing about each painting, the eye-position recording was stopped and they were asked to rate the pleasingness of the painting on a 10-point scale where 1 = not at all pleasing and 10 = very pleasing. In addition they were asked if they had ever seen the painting before to which participants reported they had not on 98% of the 120 trials.

Participants were seated 30 cm from the display monitor for an initial calibration, during which time they were not allowed to change their viewing distance, angle or head position. After a subject was calibrated, he or she was free to move side-to-side, forward and backward, and talk since their movements were tracked by a magnetic head-tracker that integrated the acquired data automatically with the eye-position data. The calibration pattern was displayed prior to each image to insure that each subject remained calibrated throughout the experiment. Data collection was initiated as soon as each image was displayed and was stopped when a subject indicated that he or she was finished examining the image and was ready to report the pleasingness rating.

Apparatus. Eye position was recorded using an ASL 4000SU Eye-Tracker with head tracking capabilities (Applied Science Labs, Bedford, MA). The recording and analysis of eye-position data have been described in detail by Nodine *et al.* (1992), but a brief overview is provided. The 4000SU eye-tracker is an infra-red 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 dwell time). The system is accurate to less than 1 degree. The raw data points are grouped into fixations and the fixations within close spatial proximity can be grouped into fixation clusters, or as in the present experiment, grid squares. Features

within the stimulus images can then be correlated spatially with the fixation data and dwell times derived. Fixation number represents the order in which the fixations were generated from the beginning of each trial, so the analysis of the data can indicate when during search various stimulus features (or grid squares, as explained below) were fixated.

Identification of the major structural aspects of the art stimuli. Five artists and five researchers engaged in experimental aesthetics investigations involving paintings identified the regions in each art stimulus that, in their view, are principal contributors to the perceptual and semantic meaning of the work. They did this by drawing circles, each with an approximate diameter of 1 in. (2.54 cm), on an $8\frac{1}{2}$ by 11 in. (21.6 cm $\times$ 27.9 cm) color copy of each painting to identify the locations. They then ranked the areas in terms of the importance of the perceptual features contained within to the structural organization and interest of the work. Agreement was 90% across the group in terms of both the locations and ranks of the areas. For purposes of analyses, we designated three areas within in each artwork that were ranked 1, 2 and 3 by at least 8 of the 10 experts as containing the work's major compositional aspects. The location of each structural aspect within the pictorial field was defined as a circular area with a diameter of 1 in. about the centroid of the cluster of circles drawn about that region by the 10 experts. The locations of the three major aspects of the structural organization of each artwork are shown on its image in Fig. 1a–h.

Results

Verbal reactions to the art stimuli. The verbal responses of each participant to each stimulus were characterized separately by the first author and his research assistant according to the reaction classification system described above. Discrepancies in classification of a response were resolved after the two sets of data were compared for inter-rater agreement which was 89%. As an example of how the system was applied, the following is one individual's response to Leger's *Woman with a Cat*; the numbers in parentheses in the sentences indicate the type of viewer reaction at that point in the time course of exploration. 'This one is much more modern (6). It's just a lot of different shapes (2). It is kind of weird (5). The face bothers me (5), with the shadows in the middle of the face (2). I like how the artist created the chair (1), the woman (1), and the cat (1) out of shapes instead of just drawing them normally (6). The woman has a book; she is reading a book (2). The yellow and black checker background (1) is out of place in the picture (6). But I think the picture is interesting (5)'.

To determine how participants initially reacted to the content and structural organization of the artworks, we examined their responses to each picture prior to 7 s of viewing time, the time by which 98% of all initial verbalizations were complete. Table 1 contains the percentages of each of the six types of reactions observed prior to the first 7 s of scanning and the chi-squared value indicates that the percentages

are significantly different. It was found that very few (6%) initial responses to the artworks referred to single compositional elements, as was the case in Experiment 1. As seen in Table 1, the majority (88%) of participants' initial reactions reflect attention to a group of pictorial elements perceived as a compositional unit, to the expressiveness of the whole composition, or to its style and form (i.e. reaction types 2, 5 and 6), as was the case following a single brief glance at the stimuli in Experiment 1. It should be noted that a significantly greater number of participants' initial reactions in Experiment 2 are of type 5 (i.e. reflecting expressiveness of the artworks) as compared to types 2 and 6, $\chi^2(2) = 9.44$, $p < 0.01$. Taken together, the results of Experiments 1 and 2 provide strong evidence that participants' initial reactions to the artworks consisted of a gist response based on the rapid processing of visual input from a large area of the visual field, as predicted by our two-stage processing model.

The frequencies of reaction types to the artworks from 7 s of viewing to the end of exploration, that is, during the cognitive stage of exploration, are presented in Table 1 along with the results of the chi-squared analysis showing that the values are significantly different. As seen in the table, more reactions (87%) were of type 2, 5 and 6 than types 1, 3 or 4, which was the case for participants' initial reactions to the artworks in this experiment and for the initial reactions of participants given a brief glance at the art stimuli in Experiment 1. Thus, participants' verbalizations in the present study did not reflect, in general, a change in the way they thought about the artworks throughout the time course of exploration.

Exploratory coverage of the pictorial fields. The fixation data were used to determine where within the limits of a 5 × 5 grid matrix superimposed over each painting, and for how long participants focused their attention within each composition during three time periods across the aesthetic experience. The purpose of the grid was to measure how eye fixations were distributed over pictorial features contained within the grid areas of each painting. Because our model of image perception during an aesthetic experience stresses the importance of the initial impression that provides a global overview of an artwork, eye fixations were analyzed separately for the first 3 s of viewing and from 3 s to 7 s of viewing to identify which pictorial elements captured the participants' attention initially. These time periods were utilized because it was found that in almost all cases participants began to speak about a composition between 2 and 3 s after it appeared on the screen and 98% of all initial verbal reactions were complete within the first 7 s of exploration. We also examined coverage between 7 s and total viewing time to determine whether interest in pictorial features which initially attracted visual attention was sustained throughout viewing, and which additional factors drew attention during the second (focal viewing) processing stage.

The grid that was superimposed over the paintings was 36.0 cm × 28.0 cm. The center of this grid was aligned to the center of the eye-movement calibration pattern. Each square of the 5 × 5 grid was 4.4 cm × 3.0 cm. This rectangle subtended a

visual angle of approximately 10 deg. horizontally at a 25 cm viewing distance or 4% of the total grid area. Using this grid, five of the paintings with horizontal orientation (the works by Bruegel, Edo Period, Giotto, Klee and Matisse) fit within the 5 × 5 grid dimensions. However, the other three paintings have a vertical orientation and the works by Leger and Marin fit just the three center columns of the 5 × 5 grid and Vermeer's composition fit the four columns to the left. Thus, for these compositions a 3 × 5 and a 4 × 5 grid, respectively, with the same dimensions as for the other paintings was used to identify regions of interest within these three paintings. By superimposing each participant's eye fixation pattern over the grid/painting image, it was possible to calculate the time spent fixating the pictorial content in each grid area within each time period of observation. The measure of time spent in each grid square was cumulative fixation duration (dwell time). Figure 1 contains the average percent dwell time across the pictorial field of each artwork at the three points of the aesthetic experience. Grid cells that do not contain dwell times in Fig. 1 are those areas of a composition that either did not receive any direct fixations, or received fixation data having less than 2 standard errors of dwell below the mean for the cell (cut off values were 238, 258 and 263 ms for the three time periods described above).

Initial coverage of the pictorial fields. Visual coverage of an artwork by a participant during one of the time periods mentioned was defined as the percentage of grid cells that contain dwell times for that artwork. For example, if a participant directly fixated pictorial elements in 5 of the 25 grid cell locations of an artwork during the first 3 s of exploration, initial coverage of that composition would be 20%.

Average percent coverage of each composition is presented in Table 2 for the first 3 s and first 7 s of exploration and the distribution of dwell times across the pictorial field of each artwork for all participants is shown in Fig. 1. As seen in Table 2, coverage of the pictorial field for the first 3 s of exploration were very similar across paintings, ranging from 20% for the Bruegel composition to 33% of the Leger and Vermeer works. Differences in coverage among the artworks shown in Table 2 are not reliable, as determined by a one-way repeated-measures analysis of variance (ANOVA), $F(7, 112) = 0.47$ ns. Given that participants began to describe their reactions to the stimuli approximately 2 s after stimulus onset, it appears that their initial responses, which were overwhelmingly holistic in nature (i.e. types 2, 5 and 6, see Table 1), were generated from direct fixation of compositional elements contained in approximately one-fourth (27%, on average) of the pictorial field.

By the time participants completed their first statements about the artworks they had expanded their coverage to 38%, on average, of the pictorial fields. As seen in Table 2 the average percent coverage for the eight artworks were relatively similar and difference among the artworks are not reliable as determined by a one-way repeated-measures ANOVA, $F(7, 112) = 1.81$ ns. Specific pictorial elements of the artworks which received attention during the first 3 s and 3 s to 7 s periods of

exploration can be seen in Fig. 1 by observing the distribution of dwells across each artwork's pictorial field.

Coverage during the second focal exploration stage of processing. Table 2 shows that participants directed their gaze to 46%, on average, of the compositions' pictorial grid fields during the entire aesthetic experience. Difference in coverage for the full viewing time did not differ reliably among the eight artworks, as determined by a one-way repeated-measures ANOVA, $F(7, 112) = 1.08$ ns. Given that average coverage of the compositions was 38% during the initial stage of exploration of the artworks, only an additional 8% of the compositions' areas received visual interest as measured by gaze dwells during the remainder of participants' interaction with the works. A 1×3 repeated measures ANOVA revealed a significant difference in average coverage among the 3 s, 7 s, and full viewing exploration period, $F(2, 21) = 9.54$, $p < 0.01$. Specifically, a Scheffe' test revealed that a significant increase in coverage occurred between 3 s and 7 s of viewing ($p < 0.01$) but the difference between coverage at 7 s and at the end of exploration was not significant ($M_s = 27\%$, 38% and 46%, respectively). Thus, participants did not examine significantly more of the compositions during the second focal processing stage as can be seen in the distribution of dwells across each artwork's pictorial field in Fig. 1.

The cells that received the greatest attention (percent dwell) when measured across all stimuli and trials are 13, 18, 17, 8 and 12 (ranks 1–5, respectively). As seen in Fig. 1, cells 8, 13 and 18 are located in the middle column of the grid and cells 12 and 17 lie to the left of the middle column (the cells are numbered 1–25 with cell 1 located in the upper left corner and cell 25 in the lower right corner counting across grid rows). The focus of attention on the contents of these five areas of the pictorial fields was consistent across the three time periods studied supporting the view that the major areas of compositional interest in the visual arts are in the center of the pictorial field.

The average total time spent by participants viewing each of the eight artworks before rating it for pleasingness is shown in Table 2. Differences between these values were not reliable when tested with a one-way repeated measures ANOVA, $F(7, 112) = 0.95$ ns. It was also observed that the median duration of participants' fixations during the global analysis and focal exploration stages were not reliably different, being 220 ms and 233 ms, respectively (Mann-Whitney U-test, $z = 1.34$ ns).

Exploration of the major structural aspects of the artworks. A participant was said to have fixated a particular major aspect of a composition if one or more fixations of his or her scanpath fell within a circular area of the pictorial field designated for that aspect by the panel of experts as explained above. The percentage of major compositional elements fixated directly during the first 3 s of viewing across the full set of 120 trials (15 participants $\times$ 8 artworks) was 6%,

27%, 50% and 17% for 0, 1, 2 or 3 elements. At 7 s of exploration the percentages were 4%, 11%, 48% and 37%, respectively. Thus, prior to verbalizing their initial reactions to the artworks, participants had encoded at least two major structural components during 67% of all trials, and this value increased to 85% by the time participants had verbalized their initial reactions to the artworks. These findings can be seen in Fig. 1 by mapping the locations of the three major aspects shown for each artwork with the distribution of dwells across the work's pictorial field at the 0–3 s and 3–7 s time periods shown in the figure. These results demonstrate that major compositional elements of the art stimuli drew viewers' interest at the earliest stage of exploration and presumably contributed to their global impressions of the art stimuli. By the time participants completed viewing each composition, they had directed their gaze to all three regions of each composition designated as containing key aspects of its structural skeleton by the panel of experts, as can readily be seen in Fig. 1.

Pleasingness ratings. Average ratings of the pleasingness of the eight art stimuli are shown in Table 2 for Experiments 1 and 2. Ratings were significantly lower, on average, for the set of stimuli following a brief glance at them in Experiment 1 than they were when viewers had unlimited time to examine the artworks in Experiment 2 ($M_s = 4.59$ *vs.* 6.17, respectively; $t(8) = 21.52$, $p < 0.01$). This was the case for 6 of the 8 artworks. Additionally, the two sets of ratings were found to be significantly correlated, $r(6) = 0.73$, $p < 0.04$. Thus, while the stimuli were evaluated as more pleasing, on average, when participants had more time to view them, differences in ratings among the artworks were relatively similar across the two viewing conditions.

Discussion

The present study adds support for the two-stage processing model of art perception proposed by the present authors (Locher, 1996; Nodine and Krupinski, 2003). Experiment 1 demonstrated that viewers are able to acquire an initial overall impression (or gist) of a painting's structural arrangement and semantic meaning on the basis of pictorial information obtained with a single brief glance. The second experiment provided insights into the nature of this initial global analysis stage of processing. Approximately 2 s after the onset of the pictorial display, participants began to describe holistic characteristics of the artworks (i.e. reaction types 2, 5 and 6). It was found that 41% of all initial responses recorded during this phase of exploration were reactions to expressive qualities of the compositions. Such qualities were derived from participants' direct fixation of the major structural aspects of the artworks that are spread across the pictorial fields. For example, it seems reasonable to assume a participant's initial assertion that the painting *The Great Wave at Kanagawa* (Fig. 1b) 'says something about the forces of nature' is based upon his or her perception of the precarious position of the boat on the left in the turbulent sea and crashing waves, the regions of the composition fixated

by this participant and all others during the initial processing stage. Moreover, the finding that the majority of participants' initial reactions to the artworks were to their style, form, and expressiveness suggests that the gist reaction to an artwork consist of more than just the perception of pictorial properties such as a composition's symmetry, balance, or complexity, structural features which research has consistently shown can be detected with little more than a 50 ms glance at an artwork (Locher and Nagy, 1996).

The fact that participants' initial reactions were based on information distributed across the pictorial field is consistent with Rasche and Koch's (2002) explanation of the nature of a gist response and the neural mechanisms responsible for it. Results of Experiment 1 demonstrate that viewers can, in fact, generate a holistic reaction to the artworks prior to a first saccade. Furthermore, Rasche and Koch (2002) assert that gist recognition is based only on a subset of an image's information. This was clearly the case in the present study; participants' initial reactions were based on direct exploration of 27%, on average, of the pictorial field during the first 3 s of the aesthetic experience. Furthermore, the regions selected for initial exploration contained major structural aspects of the artworks. While verbalizing their initial reactions, participants' scanpaths expanded significantly to include an additional 11% of the pictorial fields of the artworks.

Additional evidence that a global impression of an artwork is achieved at first glance is provided by the finding that the pleasingness ratings of the paintings obtained following a brief glance in Experiment 1 were correlated significantly with the ratings obtained following unlimited viewing in Experiment 2. This similarity suggests the evaluation of pleasingness can be made rapidly as is typically observed in the viewing behaviors of museum goers (Smith and Smith, 2001).

After the first 7 s of viewing, participants expanded their scanpaths to include an additional 8% of the compositions' pictorial fields (see Table 2), a non-significant increase. As seen in Fig. 1, they frequently returned during focal exploration to examine the three picture aspects deemed important components of the overall structural organization of the compositions and by the end of the time course of exploration, most participants had directly fixated each of these three regions in each composition. This is consistent with the visual rightness theory of picture perception (Locher, 2003). According to this theory, a 'visually right' (i.e. 'good') composition is one that has a very efficient structural organization, that is, one that is readily salient to the viewers of a work regardless of their expertise and experience with art.

The artworks that served as stimuli in the present study were created by renowned artists of recognized talent and are, therefore, presumably visually right. Thus, it is not surprising that participants' scanpaths include the major aspects of the structural organizations of the artworks during both initial and later stages of processing. They concentrated their attention in the central grid locations of the pictorial fields demonstrating 'the power of the center' of visual arts compositions (Arnheim, 1988). Grid locations in the outer regions of each painting drew little or no direct

attention as can be seen in Fig. 1. In fact, approximately 54% of the pictorial fields were not directly fixated or did not receive sustained fixation (see Table 2). Presumably, these areas serve as the backdrop for aspects of central interest in a composition. Very similar observations have been reported by other researchers (Buswell, 1935; Locher *et al.*, 1996; Nodine *et al.*, 1993). Taken together, these findings provide empirical support for the contention by art theoreticians (e.g. Arnheim, 1988; Bearden and Holty, 1969) that the most important sources of pictorial information in a painting are concentrated in the central region of a picture.

Additionally, several art stimuli used in the present study are asymmetrical with respect to their lateral organization, and lateral organization has been found to influence evaluation of paintings with right-handed individuals tending to prefer asymmetric paintings when their areas of interest are in the left pictorial field (e.g. McLaughlin *et al.*, 1983). Participants' scanpaths reflect this stimulus asymmetry in the present study. For example, the three major structural components of Giotto's *The Epiphany* (Fig. 1c) are located in the center of the left side of the composition. As one would expect, and as seen in Fig. 1, participants concentrated their attention throughout the aesthetic experience on these areas of the pictorial field, leaving much of the right field relatively unexplored. Greater concentration of exploration was also found on the left side of the works by Leger, Marin, Vermeer, and the Edo Period seascape (see Fig. 1). The present research makes clear that the influence of stimulus asymmetry on differential exploration of paintings requires systematic investigation.

In the present study, participants spent an average of 32.5 s looking at a composition before assigning it a pleasingness rating. This value is remarkably similar to the average time spent by visitors to The Metropolitan Museum of Art in New York City looking at individual paintings. In two separate studies, Smith and Smith (2001, 2003) report that visitors spent approximately 30 s, on average, with a masterpiece in the Museum's collection before moving on to view another painting. Thus, the duration of the aesthetic experience in the present study appears to be approximately as long as that for museum visitors, a finding that provides some external validity to the task employed and the data collected in the present study.

The present findings concerning the nature of the gist response also provide support for aspects of Leder *et al.*'s (2004) five-stage information processing model that underlies an aesthetic experience with visual art. According to their model, the first three processing stages involve the automatic, bottom-up perceptual analysis of pictorial features (such as symmetry, complexity, and groups of elements), of the prototypicality and familiarity of pictorial elements, and the style and content of images, respectively. The majority of participant reactions to the art works observed during the global analysis stage of exploration in the present study were to groups of compositional elements perceived as a unit, to the expressiveness of the artworks' content, and to the style and form of the compositions. These reactions correspond to the types of perceptual-cognitive analyses said to occur at stages 1 and 3 of Leder *et al.*'s (2004) model and support their view that analyses during

these initial processing stages are automatic. Thus, the initial processing stages of their model correspond to the global analysis stage of processing of our model. Furthermore, reactions to the expressiveness of both a composition's content *and* of the viewer's reaction to it are included in our fifth reaction type. Thus, some of our participants' initial reactions to the artworks would be classified as occurring in stage 4 of Leder *et al*.'s (2004) model that involves deliberate (i.e. top-down) self-related interpretations of the work. The present findings suggest that such reactions may occur much more rapidly and automatically than predicted by their model. Informal observations of visitor behavior in museums support this assertion. Visitors appear to decide rapidly not to stop and spend time with certain artworks following a brief glance at it as they proceed through a gallery, suggesting that they are able to make a rapid evaluation of a picture's content and aesthetic appeal 'at first glance'.

In conclusion, findings of the present study support the two-stage processing model of art perception described above. Furthermore, they demonstrate that knowing the way viewers visually explore *and* think about an artwork as they judge its hedonic value provides valuable insights into the perceptual–cognitive processes that underlie an aesthetic experience with visual art. To acquire a more comprehensive understanding of the nature of these processes and to subject the models of visual aesthetics discussed above to further experimental scrutiny, at least two major limitations of the present study must be addressed in future research. The first of these concerns the fact that all participants in the present study were unsophisticated in the visual arts. Eye movement studies have consistently shown that level of art sophistication is reflected in the way individuals view art. For example, both Nodine *et al*. (1993) and Zangemeister *et al*. (1995) report that professional art viewers utilize more global scanning (i.e. greater amplitude and duration of fixations) of artworks than do non-professional viewers who tend to prefer concentrating their gaze on local aspects. The second limitation of the present study is that participants performed only one evaluative task and it has been observed that the task performed on art stimuli also contributes in a differential way to a viewer's scanpath during an art experience. As one example, Zangemeister *et al*. (1995) have reported that the scanpath patterns of subjects looking at reproductions of the same abstract and realistic artworks showed different amounts of perceptual-cognitive effort when they were asked to either look at a work without further instructions, to look at it and remember its content for a recall test of specific features, or to concentrate on artistic qualities of the works. Locher *et al*. (1993) found that when individuals assessed complex abstract dot patterns that varied in symmetry for complexity or pleasingness, they concentrated their gaze in regions about the central axes of the patterns. However, when they searched the same patterns for a target, no such biased scanning was observed. Clearly, much additional research is needed before a comprehensive understanding of the nature of the perceptual–cognitive process underlying visual aesthetics is understood.

REFERENCES

Arnheim, R. (1988). *The Power of the Center*. University of California Press, Berkeley, CA, USA.

Bearden, R. and Holty, C. (1969). *The Painter's Mind*. Crown Publishers, New York, USA.

Bouleau, C. (1980). *The Painter's Secret Geometry*. Hacker Books, New York, USA.

Buswell, G. (1935). *How People Look at Pictures*. The University of Chicago Press, Chicago, IL, USA.

Ericsson, K. and Simon, H. (1993). *Protocol Analysis: Verbal Reports as Data*. The MIT Press, Cambridge, MA, USA.

Kundel, H., Nodine, C., Conant, E. and Weinstein, S. (2007). Holistic component of image perception in mammogram interpretation: Gaze-tracking study, *Radiology* **242**, 396–402.

Leder, H., Belke, B., Oeberst, A. and Augustin, D. (2004). A model of aesthetic appreciation and aesthetic judgments, *Brit. J. Psychol.* **95**, 489–508.

Lega, L., Paula-Pereira, L., Giron, S., Pastor, S. and Locher, P. (2003). A cross-cultural analysis of the visual rightness theory of picture perception, *Bull. Psychol. Arts* **4**, 86–89.

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.

Locher, P. J. (2003). An empirical investigation of the visual rightness theory of picture Perception, *Acta Psychologica* **114**, 147–164.

Locher, P. and Nagy, Y. (1996). Vision spontaneously establishes the percept of pictorial Balance, *Empirical Studies of the Arts* **14**, 17–31.

Locher, P. and Nodine, C. (1987). Symmetry catches the eye, in: *Eye movements: From physiology to cognition*, J. K. O'Regan and A. Levy-Schoen (Eds), pp. 353–361. North Holland Press, Amsterdam, The Netherlands.

Locher, P., Cavegn, D., Groner, M., Muller, P., d'Ydewalle, G. and Groner, R. (1993). The effects of stimulus symmetry and task requirements on scanning patterns, in: *Perception and Cognition*, G. d'Ydewalle and J. Van Rensbergen (Eds), pp. 5–69. Elsevier Science Publishers, Amsterdam, The Netherlands.

Locher, P., Gray. S. and Nodine, C. (1996). The structural framework of pictorial balance, *Perception* **25**, 1419–1436.

McLaughlin, J., Dean, P. and Stanley, P. (1983). Aesthetic preference in dextrals and sinistrals, *Neuropsychologia* **21**, 147–153.

Molnar, F. (1981). About the role of visual exploration in aesthetics, in: *Advances in intrinsic motivation*, H. I. Day (Ed.), pp. 385–414. Plenum Press, New York, USA.

Nodine, C. and Krupinski, E. (2003). How do viewers look at artworks? *Bull. Psychol. Arts* **4**, 65–68.

Nodine, C. and McGinnis, J. (1983). Artistic style, compositional design, and visual Scanning, *Visual Arts Res.* **12**, 1–9.

Nodine, C. F., Kundel, H. L, Toto, L. C. and Krupinski, E. A. (1992). Recording and analyzing eye-position data using a microcomputer workstation, *Behavior Research Methods, Instrumentation and Computers* **24**, 475–485.

Nodine, C., Locher, P. and Krupinski, E. (1993). The role of formal art training on the perception and aesthetic judgment of art compositions, *Leonardo* **26**, 219–227.

Rasche, C. and Koch, C. (2002). Recognizing the gist of a visual scene: Possible perceptual and neural mechanisms, *Neurocomputing* **44–46**, 979–984.

Smith, J. and Smith, L. (2001). Spending time on art, *Empirical Studies of the Arts* **19**, 229–236.

Smith, J. and Smith, L. (2003). Origins of impressionism: relating visitors' behavior to perceived learning, *Bull. Psychol. Arts* **4**, 80–85.

Zangemeister, W., Sherman, K. and Stark, L. (1995). Evidence for a global scanpath strategy in viewing abstract compared with realistic images, *Neuropsychologia* **33**, 1009–1025.