

AESTHETIC JUDGMENTS OF NOVEL GRAPHIC PATTERNS: ANALYSES OF INDIVIDUAL JUDGMENTS¹

THOMAS JACOBSEN AND LEA HÖFEL

*Institut für Allgemeine Psychologie
Universität Leipzig*

Summary.—Aesthetic judgments were investigated using a combined nomothetic and idiographic approach. Participants judged novel graphic patterns with respect to their own personal definitions of “beauty.” Judgment analysis was employed to derive individual case models of judgment strategies as well as a group model. As predicted, symmetry had the highest correlations with aesthetic judgments of beauty. Stimulus complexity was the second-highest correlate of a positive evaluation. Thus, there was agreement at the group level. The judgment analyses, however, indicated substantial individual differences. These included use of symmetry or complexity cues that were contrary to the main group use, e.g., a few participants considered nonsymmetric patterns more beautiful. These findings suggest that exclusive consideration of the group model would have leveled the individual differences and been misleading. The group model is significant; however, the individual judgment analyses represent individual patterns of judgment in a notably more accurate way.

Since Fechner’s seminal writings (1876), a number of factors have been reported which affect aesthetic judgment, including complexity (e.g., Eisenman, 1967; Berlyne, 1970), symmetry (e.g., Eisenman), novelty (e.g., Berlyne), and semantic content as opposed to formal qualities (Martindale, 1988). These factors are primarily associated with the characteristics of stimuli. In addition, emotional factors (e.g., Konecni, 1979) and interestingness (Berlyne, 1970) are factors associated with judges. Finally, social status, financial interest, and cultural background in general are known to influence preference judgments (Konecni, 1979). Symmetry and complexity were selected as independent variables for the present study because they play an outstanding role in the judgment of beauty. By varying these two formal features in newly designed stimuli, all were novel, formal, and not appealing to social status, i.e., other factors known to affect aesthetic judgment could be controlled for to achieve good experimental control.

Nomothetic and idiographic approaches need to be contrasted in research on aesthetic judgment. Most of the studies of empirical aesthetics have used a nomothetic group approach in which individual differences were

¹We are very grateful for Erich Schröger’s continued support of this project and to Volker Bosch, Ricarda Schubotz, and three anonymous reviewers for helpful comments on earlier versions of the manuscript. Address correspondence to Thomas Jacobsen, Cognitive and Biological Psychology, Institute of Experimental Psychology, University of Leipzig, Seeburgstrasse 14-20, 04103 Leipzig, Germany or e-mail (jacobsen@uni-leipzig.de).

treated as error variance. This approach might level substantial individual differences by using an averaging approach and might consequently overestimate group-level agreement. On the other hand, there are investigators (e.g., Fechner, 1876; Jacobsen & Höfel, 2001) who hold that there are substantial individual differences. Furthermore, folk psychology as well as the old sayings "beauty lies in the eye of the beholder" and "there is no accounting for taste" may lead us to assume that there are such individual differences in aesthetic judgment. Thus there is a potential area of conflict. This can be solved by taking both perspectives, nomothetic and idiographic, into account when designing a study. If results indicate sufficient agreement to support the nomothetic account exclusively, then idiographic accounts can be abandoned. In the case that there is little agreement, idiographic accounts need to be retained and the group-level account might have to be abandoned. The outcome of such a study, of course, is likely to differ contingently on the object of judgment. As a consequence, the method needs to take these factors into account.

Stimuli need to be adequate for the type of judgment studied, allowing for valid assessment of aesthetic judgments. This opens a methodological area of conflict. For instance, simple, formal stimuli are easily constructed and can be well controlled but might not elicit valid aesthetic judgments if they lie outside the realm of objects that normally elicit aesthetic judgments. On the other hand, rich stimuli may be very interesting and elicit spontaneous aesthetic judgments, but they are often unique or differ from each other on many dimensions and thus do not constitute good stimuli for a study at both the group and individual levels. An alternative is reporting data of individual judges and objects if sufficiently controlled stimuli cannot be constructed to simultaneously treat the group as well as the individual level of analysis. This approach, however, is mostly limited to the descriptive level.

The goal of the present study was the capture of aesthetic judgment strategies using visual stimuli that were adequate for the aesthetic judgment task and well-controlled. Particular consideration was given to individual differences in addition to the group-level analysis. To this end, newly designed stimuli, an economical procedure, and an approach using Social Judgment Theory were employed. A large set of formal graphic material was designed with a variation in symmetry a primary factor and complexity as a secondary factor.

Symmetry is an important visual feature in everyday life (e.g., Pashler, 1990) and a well-documented factor influencing aesthetic judgment (e.g., Eisenman, 1967; Berlyne, 1971). It was thus varied as the primary factor. Complexity was employed as the second, subordinate factor. It was operationalized by varying the number of elements in a pattern. As a consequence

of this approach, novelty as a factor was controlled for, as judges could not be familiar with the set of stimuli. Given the novel and abstract nature of the set, financial interest and social status could be neglected. In addition, an economical experimental protocol was chosen that allowed for aesthetically judging a large set of items ($k=250$ or more). Participants were asked to assign the stimuli to one of three ranking categories. In doing so, participants could judge a large set of items; such a performance would be complicated or made impossible by a rating or overall ranking procedure.

For capturing judgment strategies, the framework of the Social Judgment Theory was adopted, which defines judgment as a process that involves the integration of information from a set of cues into a judgment about some distal state of affairs (Hammond, Stewart, Brehmer, & Steinmann, 1975). The theory provides a foundation for judgment analysis, and individual as well as group models are derived. In accordance with Stewart (1988, p. 41), judgment analysis is defined as "using statistical methods to derive algebraic models of the judgment process." This is a paramorphic way of modeling, which is an input-output mapping of the judgment process under the premise that the material to be evaluated is perceptually salient. From this perspective, the perceptual process itself does not need to be modeled and only input into and output from the judgment process is considered. Here, judgment analysis was used to assess the participants' ways of making aesthetic judgments.

Based on the literature and the present stimulus design, a replication of the factors symmetry and complexity affecting aesthetic judgment was predicted (e.g., Eisenman, 1967; Berlyne, 1970). Thus, it was expected that symmetric patterns are judged to be more beautiful than asymmetric ones. Also, patterns containing more individual elements, i.e., more complex patterns, were predicted to elicit more positive judgments based on the rather low overall complexity which was due to the systematic stimulus construction. Furthermore, it was predicted that overall participants' aesthetic judgments could be predicted based on a linear combination of stimulus features in a substantial fashion following the Social Judgment Theory approach. Also, if agreement between participants permitted, a group model could be derived in the same fashion. Both of the main factors of symmetry and complexity were expected to be predictors of the group judgment strategy as well as in the individual models.

METHOD

Participants

Fifty-five young adults (15 males) participated in the experiment for course credit or partial fulfillment of course requirements. All were first- or second-year psychology students at the University of Leipzig. None of them

had received professional training in the fine arts or participated in a similar experiment before. Participants reported normal or corrected-to-normal visual acuity. One participant had to be excluded from the analysis of the experiment for noncompliance with the instructions.

Material

Two hundred fifty-two stimuli were constructed. Each consisted of a solid black circle (6.4 cm in diameter) showing a centered, quadratic, rhombic cut-out (4 cm) and an arrangement of 86 to 88 basic graphic elements (small black triangles). These were positioned within the rhombus according to a grid and resulted in a graphic pattern. The basic elements were arranged such that geometric figures like triangles, squares rhombuses, and horizontal, vertical, or oblique bars of different sizes were created. Using this collection of basic elements, the overall luminance was identical for all stimuli. Half of the patterns (130) were symmetric, i.e., a maximum of two mirroring operations giving four possible symmetry axes were permitted. The other half of the stimuli were not symmetric. Stimulus complexity was manipulated by varying the number of elements composing a pattern. Fig. 1 shows a subset of the stimuli.²

Procedure

Participants responded to 252 stimulus patterns in individualized randomized order.² They were instructed to judge each according to the picture's aesthetic value. The latter was operationalized by using the word "beauty" in the instruction, i.e., participants were instructed to make "beautiful" and "not beautiful" judgments. They were also instructed to anchor their judgment to the present stimulus set and not to take any extraneous objects or classes of objects, like paintings, works of design, or any other works of art, into consideration for their aesthetic judgments of beauty. Participants were told to take their time and spread the patterns out in front of them so that they could get a good overall impression of the stimulus set before they made their judgments. They were instructed to create three bins: one of at least 75 "beautiful" patterns, one of at least 75 "not beautiful" patterns, and a third possible category of "indifferent" pictures. This last bin could form the largest one (up to 102 stimuli) but could also contain no elements, if that was preferred. This procedure was chosen to give participants some freedom of choice while still limiting them to using all bins. The postexperimental interviews indicated that participants had no difficulties in distinguishing aesthetic and nonaesthetic patterns.

²The sets of stimuli and tabled means are on file in Document APD2002-017. For a copy, remit \$15.00 to the Archive of Psychological Data, P.O. Box 7922, Missoula, MT 59807-7922.

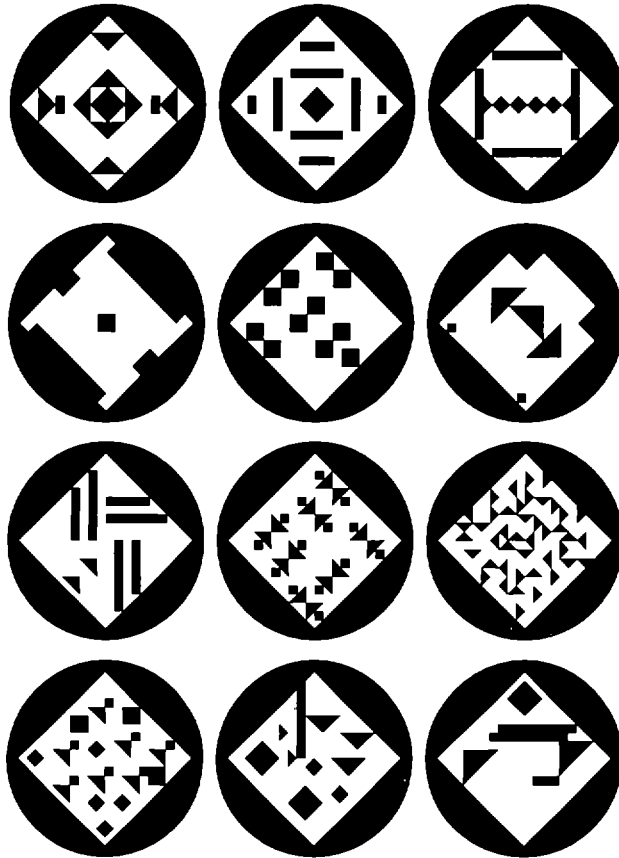


FIG. 1. Stimulus examples. The graphic patterns in rows one and two are symmetric, ranging from most beautiful to least beautiful (line by line). Patterns in rows three and four are not symmetric, also ranging from most beautiful to least beautiful.

Judgment Analysis

Participants sorted the stimuli into three bins according to the instructions. Data bins were ordered from Not Beautiful to Beautiful. For each individual judge, the normalized median of the quantile of a picture's category was assigned to the stimulus as its judgment value (McCall, 1939). For the purposes of judgment analysis of individuals these values were entered into a constrained stepwise multiple regression as the criterion along with the stimulus features as predictors (see Cooksey, 1996). The following features of the patterns were extracted: mirrored across one axis (one operation sufficient), mirrored across two axes (each one of two possible operations sufficient), mirrored over two axes (two operations required), regular com-

TABLE 1
INTERCORRELATIONS OF PREDICTORS

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1. Mirrored at one axis																				
2. Mirrored at two axes	.85																			
3. Mirrored over two axes	.82	.97																		
4. <i>Regular composition</i>	.63	.54	.56																	
5. Number of elements	-.03	.01	-.02	.16																
6. Horizontal or vertical bars	-.11	-.08	-.06	.04	.02															
7. Oblique bars	.02	.07	.06	-.15	-.28	-.15														
8. Squares	.01	.03	.01	.07	.52	-.13	-.15													
9. Rhombuses	.08	.12	.11	.04	.12	-.18	-.11	-.02												
10. Triangles	-.02	-.04	-.06	.11	.65	-.25	-.24	-.10	-.18											
11. <i>Large horizontal or vertical bars</i>	-.16	-.18	-.16	-.15	-.23	.28	-.07	-.13	-.15	-.21										
12. <i>Small horizontal or vertical bars</i>	-.06	-.02	-.01	.09	.10	.95	-.13	-.09	-.14	-.20	-.05									
13. <i>Large oblique bars</i>	-.05	.02	.01	-.20	-.37	-.17	.62	-.14	-.10	-.27	-.09	-.15								
14. <i>Small oblique bars</i>	.06	.08	.07	-.04	-.08	-.06	.80	-.09	-.07	-.10	-.02	-.06	.03							
15. <i>Large squares</i>	.15	.12	.12	.08	.08	-.15	-.10	.38	-.04	-.09	-.17	-.11	-.06	-.08						
16. <i>Small squares</i>	-.05	-.01	-.03	.04	.53	-.08	-.13	.95	-.01	-.08	-.08	-.06	-.13	-.07	.06					
17. <i>Large rhombuses</i>	.06	.07	.10	.00	-.20	-.20	-.09	-.09	.40	-.21	-.16	-.16	-.10	-.04	-.11	-.06				
18. <i>Small rhombuses</i>	.06	.10	.08	.05	.22	-.12	-.09	.01	.93	-.11	-.10	-.09	-.07	-.05	-.00	.01	.03			
19. <i>Large triangles</i>	.08	.04	.02	.01	-.11	-.24	-.18	-.20	-.20	.26	-.19	-.19	-.18	-.09	-.17	-.15	-.20	-.13		
20. <i>Small triangles</i>	-.05	-.05	-.07	.11	.71	-.17	-.18	-.03	-.11	.93	-.14	-.13	-.21	-.07	-.03	-.03	-.14	-.07	-.10	

Note.—Predictors are in boldface when the intercorrelation is above .282. These pairs of predictors were excluded from being entered into a model simultaneously.

position, number of elements, horizontal or vertical bars, large horizontal or vertical bars, small horizontal or vertical bars, oblique bars, large oblique bars, small oblique bars, squares, large squares, small squares, rhombuses, large rhombuses, small rhombuses, triangles, large triangles, and small triangles. These were considered to be perceptual cues and were used accordingly in the judgment analysis, i.e., they were introduced as predictors in the multiple regression analysis. See Table 1 for the intercorrelation matrix of all 20 features. The cue explaining most of the criterion variance was entered into the model first. Other cues were entered, providing incremental explanation of variance, if they satisfied the following constraints: there was no substantial correlation between to-be-entered cues and already-entered cues ($r < .28$), and the additional cue yielded a beta weight of $p \leq .001$ reflecting incremental explanation of variance. In addition to these paramorphic individual case models, a group model derived from the mean judgment values for each picture was computed using the same method.

RESULTS

Mean McCall-transformed ratings (McCall, 1939), standard deviations, cue values, and rank of the picture stimuli are available.² Participants reported no difficulties with performing aesthetic judgments over the material according to the instructions. Judgment analysis resulted in 53 individual paramorphic models. One participant's judgment policy could not be linearly captured. A group model was also derived. Standardized regression coefficients and multiple regression coefficients are shown in Table 2.

As predicted, symmetry was the most important stimulus feature predicting participants' aesthetic judgments. In general, participants showed agreement that the symmetric pictures were more beautiful than the others. In the group model, symmetry was also the cue explaining most of the criterion variance. Also, the individual case models had consistent and stable interindividual differences. Individual beta weights of symmetry cues ranged from .15 to .90 indicating considerable variation of cue use. Thirty-five participants used symmetry cues as the most important stimulus feature for judging Beauty in a positive direction. Twenty-two participants of this subgroup relied on symmetry cues as the sole substantial factor influencing their judgments. For them, a symmetric pattern was more beautiful. Conversely, symmetry cues were negative indicators of beauty for five participants. For four of them, symmetry accounted for the highest portion of variance. Furthermore, it was the only stimulus feature used in judgments for two of them. Symmetry was judged as not beautiful by them, a complete reversal of cue use from the overall group's pattern. That is, a minority of about 10% of the participants judged symmetric patterns to be less beautiful than non-symmetric ones, a difference that was completely leveled by the group mod-

TABLE 2
STANDARDIZED REGRESSION COEFFICIENTS (BETA VALUES) AND MULTIPLE *R*S OF INDIVIDUAL JUDGES'
PARAMORPHIC MODELS AND GROUP-MODEL BASED ON MEAN RATINGS (LAST ROW)

Subject	A	B	C	D	E	F	G	H	I	J	R
1							.57				.57
2							.88				.88
3							.90				.90
4			-.21						.40		.45
5							.34				.34
6								.53			.53
7			-.18					.59			.61
8	.23								.76		.79
9	.15						.69				.70
10									.69		.69
11							.72				.72
12									.54		.54
13	-.41									-.19	.48
14								.79			.79
15	.55									.27	.65
16							.90				.90
17								.27			.27
18									.59		.59
19	.33							.37			.49
20									.77		.77
21							-.33				.33
22	-.34							-.22			.40
23									.45		.45
24								-.28			.28
25				.15							.15
26	.21									.48	.55
27								.61			.61
28										.30	.30
29	.18							.71			.73
30								.78			.78
31							.76				.76
32								.81			.81
33							.80				.80
34	.27										.27
35										.61	.61
36	.45							.33			.55
37							.74				.74
38	-.29							-.62			.68
39							.82				.82
40	.36									-.29	.42
41		.23			.16					.62	.64

(continued on next page)

Note.—A=number of elements, B=large horizontal or vertical bars, C=large oblique bars, D=triangles, E=large triangles, F=rhombuses, G=mirrored at one axis, H=mirrored at two axes, I=mirrored over two axes, J=regular composition. Primary cue of the individual participant's judgment is set in boldface. Most important cues of the group-model are printed in italics.

TABLE 2
STANDARDIZED REGRESSION COEFFICIENTS (BETA VALUES) AND MULTIPLE RS OF INDIVIDUAL JUDGES'
PARAMORPHIC MODELS AND GROUP-MODEL BASED ON MEAN RATINGS (LAST ROW)

Subject	A	B	C	D	E	F	G	H	I	J	R
42						.28		.22			.38
43				.39		.20				.35	.57
44	.25						.27				.36
45	.18							.68			.70
46	.13						.81				.82
47							.79				.79
48	.47				.28						.52
49	.22							.67			.71
50	.23						-.22				.32
51	.31							.27			.41
52	.28										.28
53				.19					.59		.61
54											/
M	.22						.81				.83

Note.—A = number of elements, B = large horizontal or vertical bars, C = large oblique bars, D = triangles, E = large triangles, F = rhombuses, G = mirrored at one axis, H = mirrored at two axes, I = mirrored over two axes, J = regular composition. Primary cue of the individual participant's judgment is set in boldface. Most important cues of the group-model are printed in italics.

el. The nomothetic group-level approach cannot account for these judges. Moreover, the large interindividual differences in the importance a cue was given by a single judge were also leveled by the group model. Hence, it appears that detailed analysis of individual judgments provided a more adequate assessment of aesthetic judgment strategies for these stimuli.

The number of elements in a pattern, a measure of complexity, accounted for the second-most amount of variance. At the group level, pictures that contained more elements were considered more beautiful. For this, as for the symmetry cues, participants' judgments differed. While 17 participants considered stimuli with more elements to be more beautiful, three found the opposite to be the case. That is, 15% of the judges who used complexity as a cue used it in a negative direction. Beta weights (importance of the cue) varied from .13 to .55.

In addition, there were other cues that showed reverse use. For four participants, a regular composition, for instance, was the first-ranked positive predictor. Regular composition refers to patterns that are somewhat regular but need not be strictly symmetric. This feature is unrelated to complexity. Conversely, regular composition was the second ranked negative predictor for two participants. Also, compared to the features symmetry and complexity (number of elements), the shape of the elements within the frame of a pattern played a subordinate role in participants' aesthetic judgments. There were only three cases in which a shape cue explained most of the

judgmental variance. Participants 25 and 43 used triangles as the most important, positive predictor ($\beta = .15$ and $.39$), and Participant 42 used rhombuses as the most important predictor for a positive decision. Consequently, taken together, there were a considerable number of individual judges who are not well represented by the group model based on averages.

A number of stimulus features, e.g., small triangles or small squares, were not used by the participants to derive their judgments. The following predictors were not included in a model at all: horizontal or vertical bars, small horizontal or vertical bars, oblique bars, small oblique bars, squares, large squares, small squares, large rhombuses, small rhombuses, and small triangles. Participants had broad differences in linear predictability. No substantial paramorphic model was obtained for one single judge. For the remainder of participants, multiple R s ranged from $.27$ to $.90$, that is, a range of explained judgmental variance from 7 to 81%. The group model explained 74% of the variance, thus leveling the interindividual differences and hiding the huge within-subject variance. The averaging procedure treated individual difference in the aesthetic judgment as noise and thus canceled them out. In Social Judgment Theory, this difference in explained variance at the individual level is frequently interpreted as an index of strategy use (Stewart, 1988). Participants with a high linear predictability used systematic judgment strategies, while linearly unpredictable judges most likely employed highly configural cue combinations (e.g., Cooksey, 1996).

DISCUSSION

In the present study, newly designed graphic patterns were aesthetically judged by laymen. Results showed that the stimuli, albeit constructed methodically, as well as the procedure, were adequate for the aesthetic judgment task at hand. Special consideration was given to an area of conflict between the nomothetic and the idiographic approaches. Results did not show sufficient agreement between judges to support exclusively the nomothetic account and abandon the idiographic perspective altogether. Although overall agreement was found with regard to the stimulus features symmetry and complexity, the present results suggest that the idiographic perspective is at least potentially viable and needed to be retained to account for the data. As a consequence, two levels of resolution can be adopted. At a global level, the group model provides a condensed account of aesthetic judgments over the present graphic patterns in a single equation. At the local level, the results of individual judgment analyses are shown and indicate diversity between individual judges. Social Judgment Theory (Hammond, *et al.*, 1975; Stewart, 1988; Cooksey, 1996) could be employed as a useful tool in the study of aesthetic judgment of beauty.

As predicted, symmetry was the most important cue for the aesthetic

judgment of beauty of the present graphic patterns. In this way, the results converge with previous findings regarding the importance of symmetry in aesthetic judgment (Fechner, 1876; Eisenman, 1967; Berlyne, 1971). Also as predicted, complexity was the second-most important predictor of the judgments (Eisenman, 1967; Berlyne, 1970). The particular shapes of the graphic elements in the patterns, however, played a minor role. Interindividual differences reflected the fact that aesthetic judgments of beauty are not elicited in a merely stimulus-driven manner. There are a multitude of factors, including cultural and emotional aspects (e.g., Konečni, 1979), which individual judges are very likely to have encountered and coped with differently during development. Using semantically rich stimuli (cf. Martindale, 1988) instead of formal graphic patterns may result in larger interindividual differences based on the judges' respective educational backgrounds.

The group model that takes only the mean ratings into account appears more systematic and clear-cut than would be justified on the basis of the individual paramorphic models. In particular, in the group model the data of the minority of judges who used cues in a reverse fashion relative to the majority were leveled and thus excluded from consideration. Important information would be given away and the result would be inaccurate. However, the nomothetic approach does not seem to be justified in this case, and there may be situations in which a nomothetic approach by itself could be adequate.

The present procedure of computing paramorphic models was well-constrained. Cues with substantial intercorrelation could not enter the same model; the level of alpha errors when including predictors explaining incremental variance into a model was set to $p = .001$. This way the results of this study were not weakened by performing too many tests on the independent data sets of individual judges. The number of cues included into the models was well in the range previously reported (e.g., Stewart, 1988). Also, the present procedure proved to be practical and very temporally economic. Using the bin approach, participants could evaluate a large number of stimuli. This would have been much more difficult with complete ranking or a rating procedure.

In sum, the judgment analysis supported the hypothesis that symmetry and complexity are important cues for aesthetic judgments. Detailed individual models of the participants' judgment systems were derived and considerable interindividual differences were found. Some interesting evidence of reverse cue-use was obtained. Consequently, unwary use of the nomothetic approach may be improper for certain classes of objects of aesthetic judgments. In these cases, the idiographic approach, and judgment analysis in particular, can help to account for judgments of beauty.

REFERENCES

- BERLYNE, D. E. (1970) Novelty, complexity and hedonic value. *Perception & Psychophysics*, 8, 279-286.
- BERLYNE, D. E. (1971) *Aesthetics and psychobiology*. New York: Appleton-Century-Crofts.
- COOKSEY, R. W. (1996) *Judgment analysis: theory, methods, and applications*. San Diego, CA: Academic Press.
- EISENMAN, R. (1967) Complexity-simplicity: I. Preference for symmetry and rejection of complexity. *Psychonomic Science*, 8, 169-170.
- FECHNER, G. T. (1876) *Vorschule der Aesthetik*. Leipzig: Breitkopf & Härtel.
- HAMMOND, K. R., STEWART, T. R., BREHMER, B., & STEINMANN, D. O. (1975) Social judgment theory. In M. Kaplan & S. Schwartz (Eds.), *Human judgment and decision processes*. New York: Academic Press.
- JACOBSEN, T., & HÖFEL, L. (2001) Aesthetics electrified: an analysis of descriptive and symmetry and evaluative aesthetic judgment processes using event-related brain potentials. *Empirical Studies of the Arts*, 19, 177-190.
- KONECNI, V. J. (1979) Determinants of aesthetic preference and effects of exposure to aesthetic stimuli: social, emotional and cognitive factors. *Progress in Experimental Personality Research*, 9, 149-197.
- MARTINDALE, C. (1988) Aesthetics, psychobiology, and cognition. In F. H. Farley & R. W. Neperud (Eds.), *The foundation of aesthetics, art and art education*. New York: Praeger. Pp. 7-42.
- MCCALL, W. A. (1939) *Measurement*. New York: Macmillan.
- PASHLER, H. (1990) Coordinate frame for symmetry detection and object recognition. *Journal of Experimental Psychology: Human Perception & Performance*, 16, 150-163.
- STEWART, T. R. (1988) Judgment analysis: procedures. In B. Brehmer & C. R. B. Joyce (Eds.), *Human judgment: the SJT view*. Amsterdam: North Holland. Pp. 41-74.

Accepted September 2, 2002.