

---

## RESEARCH NOTE

---

### The Visual Aesthetic Sensitivity Test Applied to Portuguese Children and Fine Arts Students

**J. P. Frois**

*International Association of Empirical Aesthetics*

**Hans J. Eysenck**

*Institute of Psychiatry  
University of London*

---

**ABSTRACT:** *The Visual Aesthetic Sensitivity Test (VAST; Goetz, 1985) was administered to 722 male and female children between the ages of 10 and 15 and 58 fine arts students. Also administered to the children only were Raven's Progressive Matrices and the Eysenck Personality Questionnaire (H. J. Eysenck & S. B. G. Eysenck, 1975). VAST scores increased with age up to 14, but not beyond; sex had little influence. IQ accounted for only 10% of the VAST variance. Personality failed to correlate with VAST scores, and artistic training had no effect on VAST scores. The test scores were reasonably reliable ( $r = .70$ ), but skewed to the right, with a long tail of low scores. Difficulty levels were similar to those observed in other countries and cultures.*

The experimental study of aesthetic preferences and experiences began with the early work of Fechner (1876) and Baltalon (1900). The major aspect of this work was the introduction of the analysis of preference judgments as tools for the attempt to make the study of aesthetics scientific, discover lawful

relation in this field (H. J. Eysenck, 1964, 1981, 1983a, 1983b, 1988), and suggest explanatory hypotheses (Berlyne, 1971). Attempts to make concepts like "beauty" and "good taste" acquire an empirical basis led to the discovery of two major determinants of preference judgments in the visual field. The first of these, good taste was based on the empirical discovery that judges tended to agree on the ranking of aesthetic objects and that those who agreed best with the average rankings of one class of objects tended to agree best with the average ranking of other classes of objects (H. J. Eysenck, 1940). Classes of objects so studied ranged from

---

We are indebted to the Department of Fine Arts of the Calouste Gulbenkian Foundation for sponsoring this study.

We thank Arlette Verhaeghe for help in the organization and discussion of the results, Henriqueta Ruivo for the collection of the data, and Nitah Camotim for translating the original draft of the article.

Correspondence and requests for reprints should be sent to Hans J. Eysenck, Institute of Psychiatry, University of London, DeCrespigny Park, Denmark Hill, London, SE5 8AF, England.

simple colors (H. J. Eysenck, 1941a) and shapes (H. J. Eysenck, 1941b, 1942, 1973) to paintings and sculptures (H. J. Eysenck, 1940). The second factor, orthogonal to the first, was a preference for *complexity* as opposed to *simplicity* (C), in which complexity is defined in terms of the collative properties of the stimulus materials (Berlyne, 1971; H. J. Eysenck, 1941c). Welsh (1975) used this conception to create a test of creativity, the Barron–Welsh Art Scale (H. J. Eysenck & Castle, 1970).

There have been many attempts to measure the good taste factor, such as the Meier (1942) Art Tests and the Graves (1943) Design Judgment Test; all suffer from the same fault, namely the poor artistic quality of the drawings used. There are also psychometric faults, such as lack of homogeneity, which disqualify these tests from being seriously considered as proper tools for measurement (H. J. Eysenck, 1967, 1970; H. J. Eysenck & Castle, 1970, 1971; K. O. Goetz & K. Goetz, 1974).

To remedy this situation, a new test—the Visual Aesthetic Sensitivity Test (VAST; K. O. Goetz, 1985)—was introduced that was made up of artistically meaningful drawings made by a well-known German painter, K. O. Goetz (1985). The test has been described in detail elsewhere (H. J. Eysenck, 1983a, 1983b). It consists of a series of 50 pairs of nonrepresentative drawings; in each pair the subject has to express an opinion as to which is the better, more harmonious design. The artist drew the better designs first and then changed it by incorporating faults and errors into it. The validity of the judgment was tested by having eight expert judges (artists and critics) make preference judgments and only accepting pairs of designs where agreement among the judges was unanimous. Another proof of the validity of the original

judgment is that when groups of subjects are tested, the majority judgment always agrees with the keying of the items. This agreement between experts and unsophisticated members of the public is difficult to explain in terms other than some universal sense of aesthetic appreciation possessed in different degree by everybody—few scores below chance have been observed. Figure 1 shows three typical items from the test, ranging from *easy* through *average* to *difficult*. The difficulty level of items is established in terms of the percentage of right responses; the more subjects give the right answer, the easier the item. Different groups of subjects, differing in age, sex, artistic training, cultural



Figure 1. Easy, medium, and difficult items on the Visual Aesthetic Sensitivity Test.

background, and race have always produced very similar difficulty levels for the items.

It should not be thought that experts and untrained subjects agree because both are exposed to the same cultural environment; several studies have shown essential similarities between the test performances of British, German, American, Japanese, and Chinese groups (Chan, H. J. Eysenck, & K. O. Goetz, 1980; H. J. Eysenck, K. O. Goetz, Han Yee Long, Nias, & Ross, 1984; Iwawaki, H. J. Eysenck, & K. O. Goetz, 1975). Artistic training also has not appeared to have much influence on scores. Other studies of the VAST have demonstrated adequate psychometric properties for group testing, absence of sex differences, and small correlations with IQ or personality (K. O. Goetz, Borisy, Lynn, & H. J. Eysenck, 1979). There are two versions of the test, on slides and in printed form; the latter seems to have higher reliability. There are also several forms of the test, varying in number of items; the present version has 50 items and is (provisionally) the final version.

The aim of this study was to obtain psychometric information on the VAST from groups of children and adults, analyze the influence of age, sex, personality, and intelligence, and investigate the influence of artistic training.

## Method

### Subjects

The study was carried out with a sample of 780 subjects: 341 boys (in years, 63 were age 10; 54 age 11; 49 age 12; 68 age 13; 55 age 14; and 52 age 15), 381 girls (in years, 51 were age 10; 51 age 11; 73 age 12; 79 age 13; 69 age 14; and 58 age 15), and a group

of 58 adults consisting of 20 men and 38 women with an average age of 21 years. The sample of children originated from the same geographical region, and in it were represented diverse socioeconomic levels. The adults were students of the third and fourth years of the courses of Sculpture and Painting of the Lisbon Fine Arts School.

### Intelligence and Personality

As a measure of general intelligence, Raven's Progressive Matrices Test was used. As a measure of the major dimensions of personality (P = Psychoticism, E = Extraversion, N = Neuroticism, L = Dissimulation), the Portuguese version of the Eysenck Personality Questionnaire (EPQ; H. J. Eysenck & S. B. G. Eysenck, 1975; Fonseca & S. B. G. Eysenck, 1989) was used.

### Visual Aesthetic Sensitivity

We used the final edition of the VAST, already described. Small groups of 20 or so subjects were tested, using the printed version of the test. This constituted the first testing session. In the second, the EPQ was administered and a week later the matrices. The instructions for the VAST were as follows:

This booklet contains a series of 50 pairs of pictures. Each pair consists of two similar pictures one of which is superior from the point of view of design; it is more harmonious, better balanced and better adapted in the way the elements are ordered, and in the way the items are drawn. Look carefully at the picture, and you will see that in the comparison the worse picture contains small "faults" or "errors" which destroy the balance of the picture. These judgments were made unanimously by a group

of painters and graphic artists for all 50 pairs of pictures.

At the beginning of the booklet are given three examples (pairs A, B and C) which illustrate these points. A cross is affixed to the more balanced, better formulated picture. In the case of picture A, the picture marked with a cross is better because the contour is more even, and the central point is located right in the middle. In the less good picture there are breaks in the contour in two places, and the circle is badly positioned.

In the case of B, in the better drawing the three polygon figures present a well-ordered and circumscribed design, whereas in the worse drawing they fall apart altogether.

As regards Figure C, the rhythmic movement can be seen clearly in the better figure, whereas this is not at all clear in the worse figure, where the small pointed areas constitute a disturbance.

Your task will be to discover, as far as the fifty remaining pairs of pictures are concerned, which in each pair is the better design. Sometimes this will be on the left side, sometimes on the right. Look carefully and take your time; the correct answer is not always obvious! When you have made a decision, write a capital R (for right) or a capital L (for left) in the space after the number of the picture pair on the score sheet you have been given. Do not leave out any of the pairs, even though you may find difficulties in coming to a decision.

Do not form your judgment according to which of two pictures you like better. This is not the point, because sometimes a picture that is less well designed may please you better, because you find it personally more interesting. Your task is to discover which picture has been better designed. Note that the problems differ in difficulty, but that the more difficult ones are not necessarily to be found near the end; easy and difficult items are intermingled on a fairly random basis.

For the other tests, testing and instructions were as suggested in the test material. Only the children received the EPQ and the matrices; the Fine Arts students only worked on the VAST.

## Results

The alpha reliabilities of the test are given in Table 1 for the various groups. Those for the adults are satisfactory (.72 for both males and females); for the children they are variable, ranging from .51 to .70 for the combined sex group; clearly the numbers are too small within each group to make the differences meaningful. There is no clear trend for age or sex. Scores are skewed to the right, with a long tail sloping toward low scores; however, the skew is not extensive enough to invalidate the use of standard psychometric procedures.

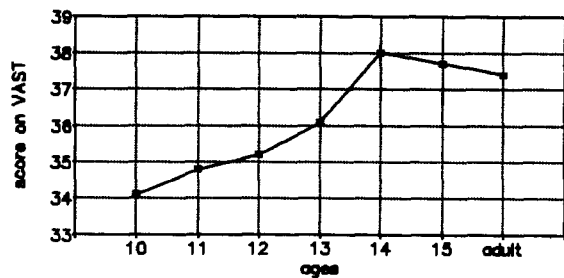
Table 2 shows the means and standard deviations for the various groups and the significance of the differences. Four of these are insignificant, but for ages 11, 12, and 13, females have significantly higher scores. There is a possibility that this may be connected with puberty, but only replication can tell whether these rather small differences have any substantive importance and require explanation. Analysis of variance for the total sample shows that the principal effects of the age and sex factors are highly significant, with  $p$  values of .01 in both cases. The interaction between sex and age is also significant ( $p < .025$ ) but barely so; this interaction mirrors the pattern of significance levels shown in Table 2.

**Table 1.** Alpha Reliabilities of the Visual Aesthetic Sensitivity Test for Males and Females of Different Age Groups

Age (Years)	Total	Males	Females
10	.509	.514	.498
11	.620	.530	.575
12	.669	.710	.600
13	.698	.685	.676
14	.557	.565	.558
15	.559	.584	.493
Adults	.725	.722	.734

**Table 2.** Differences in Means and Standard Deviations for the Visual Aesthetic Sensitivity Test at Different Age Levels

Age (Years)	Boys and Girls			Only Boys			Only Girls			F
	n	M	SD	n	M	SD	n	M	SD	
10	114	34.1	4.4	63	33.8	4.5	51	34.7	4.3	$F(1, 113) = 1.21, ns$
11	105	34.8	5.0	54	32.8	4.7	51	36.9	4.5	$F(1, 104) = 20.6, p < .01$
12	122	35.2	5.3	49	33.8	5.9	73	36.1	4.8	$F(1, 121) = 5.73, p < .05$
13	147	36.1	5.4	68	34.8	5.5	79	37.3	5.1	$F(1, 146) = 8.26, p < .01$
14	124	38.0	4.3	55	38.1	4.4	69	38.0	4.3	$F(1, 123) = 0.014, ns$
15	110	37.7	4.3	49	37.0	4.1	58	38.4	4.0	$F(1, 106) = 2.99, ns$
Adults	58	37.4	5.5	20	37.4	5.5	38	37.4	5.6	$F(1, 58) = 0.002, ns$



**Figure 2.** Scores on the Visual Aesthetic Sensitivity Test (VAST) as a function of age. Table 3 shows the correlations between the VAST and intelligence test scores for the children.

As far as age is concerned, Figure 2 illustrates the results. Clearly scores on the VAST increase with age up to 14, after which point there is no change. The difference between the 14- and 15-year-old groups, on the one hand, and the adults on the other, are not statistically significant.

We now turn to a discussion of the relation between the VAST and intelligence. Table 3 shows the correlations between the two sets of scores for the children. All the product-moment correlations are positive, although not all are significant; they average around .35, and none depart significantly from this mean. Roughly 10% of the total variance on the VAST is accounted for by intellectual factors in these children, leaving 90% to be specific to the VAST. This is in general agreement with previous investigations.

**Table 3.** Product-Moment Correlations Between the Visual Aesthetic Sensitivity Test and Raven's Matrices for Different Age Groups

Age	Boys	n	Girls	n	Boys/Girls
10	.33*	61	.26	50	.31**
11	.28	49	.29	49	.36**
12	.21	40	.45**	55	.31*
13	.34	39	.07	51	.20
14	.21	36	.25	50	.24
15	.13	40	.50**	51	.33*
All	.32**		.35**		.36**

**Table 4.** Product-Moment Correlation Between the Visual Aesthetic Sensitivity Test and Eysenck Personality Questionnaire Scales

Group	P	E	N	L
Males	-.035	.096	-.062	-.1705*
Females	-.019	.090	.088	-.2710**
Males and Females	-.068	.051	.044	-.2240**

Note: P = Psychoticism, E = Extraversion, N = Neuroticism, L = Dissimulation.

\* $p = .01$ . \*\* $p = .001$ .

Personality and the VAST show no significant correlations (Table 4). The only significant correlations are between the VAST and the L (Dissimulation) scale, and these are negative. L scores are often interpreted as indicative of conformity when there is no special incentive to dissimulate, so the data may suggest the more conformist children do marginally worse on the VAST. This finding

may or may not be worth following up; the percentage of variance explained is quite small. For the different groups, only one correlation is significant by itself (11-year-old girls,  $r = -.49$ ); 9 out of 12 are negative.

Difficulty levels showed no obvious differences between groups related to age or sex, except that the absolute level was of course lower for the 14- and 15-year-olds and the adults. They were in good accord with difficulty levels of other groups studied previously (H. J. Eysenck et al., 1984; K. O. Goetz et al., 1979); clearly there is considerable lawfulness about the VAST.

### Discussion

The major results of this study are in good agreement with those of (usually much smaller) other studies carried out in various countries (e.g., Chan et al., 1980; H. J. Eysenck et al., 1984; H. J. Eysenck & Iwawaki, 1971; K. O. Goetz et al., 1979; Iwawaki et al., 1975). It may be useful to summarize them here.

1. Performance on the VAST is reasonably reliable, with Cronbach's alpha levels around .70, somewhat lower for younger children.

2. Performance on the VAST increases with age up to 14 or so, and then remains steady into adulthood.

3. Performance on the VAST is largely independent of sex, although between 11 and 13, girls appear to do slightly better than the boys.

4. Performance on the VAST correlates with IQ at all ages at about  $r = .35$ , IQ thus accounting for some 10% of the VAST variance.

5. Personality does not seem to predict VAST performance, all correlations (with

the possible exceptions of the Lie scale) being insignificant.

6. Artistic training seemed to have no effect on VAST scores, as the artistically trained adults failed to do better than the (untrained) 14- and 15-year-old children. This finding strengthens the belief that genetic factors may be operating here, a belief based on the cultural independence of the VAST. Empirical support for this belief is being sought.

Clearly, the test could be made more reliable by extending it to 100 items, but boredom resulting from extended testing may reduce the gains so achieved. Choice of one from four, rather than from two items may also reduce the influence of chance choices. There are clearly ways in which the test could be improved, but it is perhaps already sufficiently well designed psychometrically to be used in experimental investigations of the intriguing and controversial concept of good taste. Clearly, the VAST is measuring something other than IQ or personality; equally clearly this "something" is related to visual aesthetic sensitivity. Much research will be required to elucidate the precise nature of this something; but it does seem to link up with the subject matter of earlier studies (H. J. Eysenck, 1940, 1941b, 1942). It is impossible to predict the extent to which the theories put forward may explain the results achieved with this test built to investigate prediction made from those theories; so far results have been in line with expectation.

### References

- Baltalon, T. P. (1900). *Nabliudenia i Opyty po Estetike Zritelnikh Vospriatii* [Observation and experimentation in visual aesthetic perception].



Meier, N. C. (1942). *The Meier Art Tests*. Chicago: Bureau of Education and Research and Service.

Welsh, G. S. (1975). *Creativity and intelligence: A personality approach*. Chapel Hill, NC: Institute for Research in Social Science.