THE GENERAL FACTOR IN AESTHETIC JUDGEMENTS[1]

By H. J. EYSENCK

From the Psychological Laboratory, University College, London

I. Previous investigations

Several workers have recently been investigating the existence and nature of a 'general factor' in aesthetic appreciation. Dr Dewar has shown that both laymen and experts tend to put pictures of widely different merit in much the same order of liking; if their various rankings are correlated, and the resulting table factorized, we find that a large part of the variance is accounted for by a general factor(1).

Other investigators have shown reason to assume the existence of a general factor which would be operative not only within one test, but between tests of different kinds of aesthetic appreciation (even after the influence of intelligence had been eliminated). They found that "the capacity for literary appreciation correlates... with pictorial and musical appreciation" ((2), p. 283).

The general factor to be discussed here lies between these two factors, being more general than that discovered by Dr Dewar, but less extensive than that discovered by Williams, Winter & Woods. Admitting that there is a general factor appearing within any one test of visual appreciation, we ask if that factor would extend to all other tests of visual appreciation as well; or whether it would be specific to the one test in which it was observed.

Prof. Burt put the same question in the following form: "If we could brush aside all irrelevant associations, and take a completely detached view... would there be any solid grounds for preference left?" ((3), p. 289). Researches such as those mentioned above may seem to answer this question *a fortiori*, because if we admit the existence of an intersensory

[1] A paper read before the British Psychological Society, 1 April, 1939.

general factor, extending over literary and musical as well as pictorial material, the existence of a general factor for visual appreciation would seem almost inevitable. This argument is strengthened by the fact that Dr Dewar found that several different tests of visual aesthetic appreciation correlated together, even after the influence of intelligence had been eliminated.

And yet we feel unable to accept these results as finally settling the question raised by Prof. Burt. It would seem that the condition mentioned by him had not been fulfilled in these researches, viz. that all irrelevant associations should be brushed aside. The 'irrelevant associations', as they appear in the researches mentioned, are mainly of three kinds:

First, there is the influence of that complex of factors which we may call 'civilization'. Included here are teaching, tradition, general knowledge and all the thousand and one things that make up our cultural background. Dr Dewar's pictures, for instance, included "reproductions from classical masters, second-rate pictures by second-rate painters, every variety and type down to the crudest and most flashy birthday card" ((3), p. 290). Now it is easy to see that the influences mentioned above might exert an overwhelming force in favour of the 'socially accepted' pictures, such as the old masters, and against the 'birthday cards'.[1] Dr Dewar herself reports the influence of a factor of this nature (4).

Two minor factors working in the same direction we might call excellence of technique and familiarity. So far as technical excellence is concerned, the old masters in most cases are superior to the artists who draw the birthday-card type of picture, in the sense of being able to reproduce external objects more accurately. But preference on these grounds is not what we are looking for; we may prefer one photograph to another, not because it is superior aesthetically, but because the other one is technically imperfect—it may be out of focus, or underexposed. This is clearly an extraneous factor, whose influence should be minimized.

As regards familiarity, the attachment one forms so readily to things one knows well has often been remarked upon. When we encounter a picture we have seen before, or a picture painted in a familiar manner, the very fact that we have seen this picture, or a similar one, before disposes us more favourably towards it. The reaction against any new school of art is an eloquent witness to this contention.[2] At first, the un-

[1] This would be true primarily with regard to the 'experts' from whose rankings the correlations are taken on which Dr Dewar bases her general factor.

[2] There is an interesting discussion of this point in Koffka's *Principles of Gestalt Psychology*, p. 347.

familiar is disliked; as it becomes more familiar, liking increases. This has been proved experimentally by Moore (5). (It is not maintained of course that this effect is invariable and certain in all circumstances.)

If these criticisms have any validity, we should be able to derive certain principles from them as regards selection of pictures free from these 'irrelevant associations'. The pictures used in experiments of this kind should fulfil the following conditions: No tradition or teaching should point to one of them as superior to the others; so far as execution is concerned, they should be roughly of the same degree of excellence; and they should all be equally unknown to the subjects. Only if these conditions are fulfilled can we regard any test as genuinely testing the question at issue.

II. General plan of the present research

Within reason, the pictures selected for the present research comply with these three conditions. Eighteen sets of pictures were used altogether, each containing over fifteen pictures on the average. Nearly all of these were completely new to the subjects; even where they were not, no authority could point to one of them as superior to the others. The standard of technical excellence was uniform throughout each test. The sets used are described below:

Description of tests used in research on aesthetic appreciation

(1) *Portraits.* Thirty-two reproductions of comparatively unknown paintings by modern painters and old masters.

(2) *Emperors.* Fifteen photographs of statues of Roman emperors, taken from the British Museum.

(3) *Book-bindings.* Fifteen coloured reproductions of book-bindings, from the British Museum.

(4) *Claude Lorrain.* Twenty reproductions of pencil drawings by Claude Lorrain.

(5) *Blotters.* Twelve coloured photographs of 'Bathing Girls', used for advertising blotters.

(6) *Vases.* Sixteen coloured photographs of vases from the British Museum.

(7) *Masks.* Eight coloured reproductions of Malayan devil masks.

(8) *Japanese paintings.* Fifteen uncoloured Japanese paintings.

(9) *Ships.* Nine photographs of modern steamships.

(10) *Landscapes.* Twelve coloured landscape paintings by Japanese artists.

(11) *Sets.* Ordering the seventeen sets of pictures in order of liking.

(12) *Embroidery.* Twelve coloured reproductions of pieces of embroidery, from the Prince Albert and Victoria Museum.

(13) *Curves.* Twelve curves of mathematical functions, drawn in ink.

(14) *Statues.* Thirty-two reproductions of modern statues.

(15) *Flowers.* Fifteen coloured photographs of various flowers.
(16) *Pottery.* Twelve coloured reproductions of pottery from the East.
(17) *Plate.* Fifteen reproductions of pieces of silver plate.
(18) *Clocks.* Twelve reproductions of medieval clocks.

The eighteen subjects taking part in the experiment were drawn from various walks of life.[1] They included bank clerks, shorthand typists, painters, students, teachers, psychologists, and a Professor of Aesthetics. The number of subjects to be used was determined by the intercorrelations of their rankings; the object being to obtain an average ranking which for all practical purposes was equivalent to the 'true order', i.e. to the average order of the whole population. I have shown elsewhere on what reasoning the formulae used in this connexion are based, and that they can be applied to material of this kind (6). The average correlation between the average rankings and the 'true order' is 0·87, as is shown in Table I.

Table I

Tests	r_{ag}	r_{as}	$\bar{r}_{k\bar{k}'}$	$\bar{r}_{kk'}$	$r_{\bar{k}g}$	N
(1) Portraits	0·811	−0·005	0·40	0·16	0·88	32
(2) Emperors	0·810	0·232	0·51	0·26	0·93	15
(3) Book-bindings	0·661	−0·524	0·27	0·07	0·76	15
(4) Claude Lorrain	0·577	−0·277	0·40	0·16	0·88	20
(5) Blotters	0·571	0·480	0·40	0·16	0·88	12
(6) Vases	0·541	0·191	0·40	0·16	0·88	16
(7) Masks	0·478	−0·545	0·26	0·07	0·76	8
(8) Japanese paintings	0·461	0·074	0·39	0·15	0·87	15
(9) Ships	0·334	0·073	0·35	0·12	0·84	9
(10) Landscapes	0·327	0·163	0·49	0·24	0·92	12
(11) Sets	0·309	0·169	0·66	0·44	0·96	17
(12) Embroidery	0·244	0·152	0·33	0·11	0·83	12
(13) Curves	0·228	−0·503	0·52	0·27	0·94	12
(14) Statues	0·202	0·592	0·35	0·12	0·84	32
(15) Flowers	0·167	0·639	0·32	0·10	0·82	15
(16) Pottery	0·058	0·035	0·47	0·22	0·91	12
(17) Plate	−0·060	0·484	0·42	0·18	0·89	15
(18) Clocks	−0·287	0·397	0·30	0·09	0·78	12
Variance	0·206	0·137				
Average			0·40		0·87	15·6

r_{ag} = factor saturation, general factor.
r_{as} = factor saturation, bipolar factor.
$\bar{r}_{k\bar{k}'}$ = average correlation of individual rankings with average order.
$\bar{r}_{kk'}$ = average intercorrelations.
$r_{\bar{k}g}$ = correlation of average order with true order.

Having asked the subjects to rank the pictures in each of the eighteen sets in order of liking, one of two alternative courses was adopted to deal with these 324 rankings. The first was to correlate the rankings, factorize

[1] I should like to take this opportunity of expressing my gratitude to these subjects, all of whom devoted to the tests of this research a great deal of time and energy—four sessions of an hour each being the minimum time for testing.

the resulting table, and thus obtain weighted saturation coefficients. This was done when there was reason to suspect that bipolar, secondary factors played an important part in the choice, as for example in the 'Portraits' test, where the variance of the second factor is as high as 15%.[1]

In the majority of cases, however, a rough preliminary analysis showed no trace of a significant second factor, and then a simpler method was adopted.[2] The average of the rankings of each set was calculated, and each person's ranking correlated with this average, thus giving what might be called 'unweighted saturations'. Sixteen tests were found to contain no important secondary factors; only the 'Portraits' and the 'Statues' tests required to be fully analysed. The average of all the 324 saturations was almost 0·40; that is to say, on the average the subjects agreed to that extent with the average order. As each person's ranking contributes to the average order, however, the average saturation would be reduced to 0·34 (using a formula given by Garrett ((7), p. 460)).

The fact that these saturations are significantly positive[3] proves the existence of a general factor for each of the eighteen tests; it will be seen that so far our results bear out Dr Dewar's findings. But the factors are still not general in the sense we mentioned at the beginning; they may be specific to each of the tests, and not general from test to test. (As, for instance, a subject's score in an intelligence test may be specific to that test, and not allow any prediction as to his score in other tests.) To determine whether there is or is not a general factor of this kind, we may regard each person's (weighted and unweighted) saturations as his scores in the tests, and correlate the tests with one another. (Our procedure follows that of the intelligence tester, who correlates several tests and deduces factors from the pattern of their intercorrelations.)

The table of correlations can be factorized by means of one of the usual formulae; the factors running through the tests can be extracted, and an interpretation of these factors attempted on the basis of the experimental material. This method of dealing with the scores gives us saturations for the tests, showing how far they are saturated with the factors involved.

[1] The influence of bipolar or 'type' factors in aesthetic judgements, their nature and relation to temperamental qualities, will be dealt with in a later article.

[2] I am indebted to Prof. Burt for suggesting this method. Also, I should like to express my gratitude to him for reading this manuscript, and making many valuable suggestions.

[3] Eighty-one out of 324 saturations were 0·64 or over. By chance, we should expect only 1 in 100 to be as large as 0·64, according to Fisher's method.

III. RESULTS AND CONCLUSIONS

In Table I are shown the results of a factor analysis of the correlations between the eighteen tests used in this research. Fig. 1 gives a geometrical picture of the saturation pattern of the two factors responsible for all the correlations, within the limits of the probable error. It will be seen that we are dealing, first, with a general factor accounting for 20·6% of the variance, and secondly, with a bipolar factor, accounting

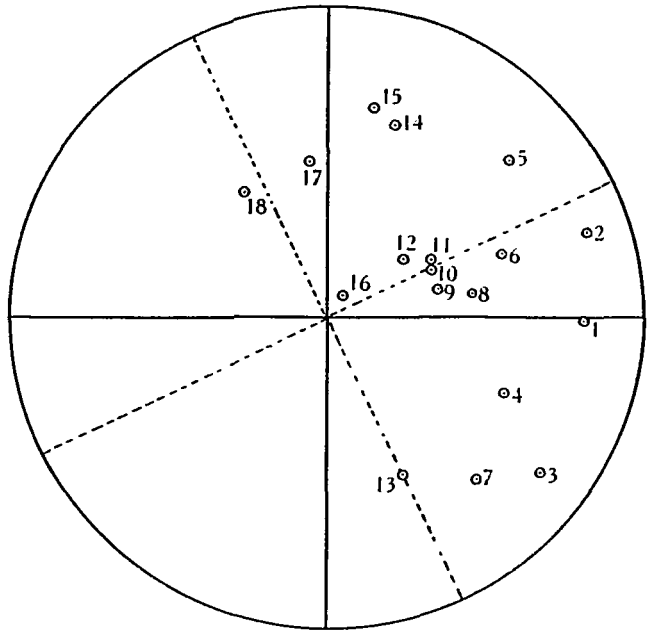

Fig. 1.

for 13·7%. Two tests have negative saturations for the 'general' factor, but these are extremely low and can hardly affect the generality of that factor. (The correlations between the two tests whose saturations are negative is only 0·089; this shows that we are not dealing with a genuine bipolar factor.)

In accordance with Thurstone's principles(8), we should rotate the axes in Fig. 1 till no negative values are left for the general factor, and the number of zero saturations is maximal. This rotation is shown graphically in Fig. 1, but as can be seen this rotation would have little effect on the main results. In certain situations, rotation has the effect

of eliminating the general factor altogether (9, 10), but this would not be the case here. The only effect would be a slight shift in the saturations for both the general and the bipolar factor.

Other results recorded in Table I are explained there. These values were calculated by formulae given by Burt (11); I have shown elsewhere that these formulae can legitimately be applied to data such as we are dealing with (6).

The nature of the general factor which we have found has already been indicated to some extent at the beginning of this article. It may be defined as the general, objective factor of aesthetic appreciation. We call it 'general' because it covers a large number of, probably all, pictorial tests, just as Spearman's '*g*' covers a large number of, if not all, cognitive tests. We call it 'objective' because, as has been shown, it is largely independent of individual taste. The word 'objective' has been used in this sense by Burt (3), Wells (12), and others; it is not here taken to imply anything with regard to the independent existence of the object. This is a philosophical question which cannot be dealt with in a psychological article.

But a few words may not be out of place with regard to the common-sense interpretation of the general factor. As applied to persons, this factor is the core of reality behind what is generally called 'good taste' (whence it has been suggested that this factor should be called '*T*'); as applied to pictures, it accounts for what we call 'beauty'; and finally, as applied to tests, it shows how good a measure of '*T*' each test provides. The 'taste' of a person can be measured by 'correlating persons'; the goodness of a test by 'correlating tests'; the relative 'beauty' of the pictures by calculating the (weighted or unweighted) average of the scores given to it by the subjects.

If the nature of the general factor is thus fairly clear, the nature of the bipolar factor is less obvious. It is possible to define its meaning by examining those tests which are most highly saturated with it. Curves, masks, and book-bindings have high negative saturations; flowers, statues, plate, clocks, and blotters have high positive saturations. An inspection of the pictures in these sets suggests immediately a dichotomy well known to aestheticians—that between 'formal' and 'representative' art.

Curves, book-bindings, and devil masks are formal in the sense that they are concerned mostly with relations of line, colour, or volume; the flowers, statues, blotters tests, on the other hand, are not so much concerned with these relations, but rather with the representation of external objects, e.g. human bodies, flowers, etc. This interpretation is supported by the introspective reports of the subjects.

Thus we have obtained experimentally a confirmation of a theory often advanced by aestheticians; Kant's distinction between 'free' and 'dependent' beauty(13), Porena's distinction between 'il bello immediato' and 'il bello di rapporto'(14), and the theories associated with the names of R. Fry(15) and C. Bell(16), all seem to emphasize the same dichotomy between what we called 'formal' and 'representative' art.

Both the general and the bipolar factor are more comprehensive than has been indicated so far. As this article is concerned mainly with the general factor, a discussion of the formal-representative factor must be left for a subsequent article. As regards the general factor, a number of the subjects whose saturations for this factor were known took part in further experiments involving the ranking of thirty-one odoriferous substances, of ten colours, and of sixty-four polygonal figures. In each case a general factor appeared when the intercorrelations of these rankings were calculated. When the saturations for these observers were correlated with their saturations for the 'T' factor, the following results were obtained:

Odours and 'T' $= 0{\cdot}75 \pm 0{\cdot}17$
Colours and 'T' $= 0{\cdot}53 \pm 0{\cdot}22$
Polygons and 'T' $= 0{\cdot}68 \pm 0{\cdot}15$

(The number of subjects in these three experiments was eight, twelve and fourteen respectively. By Fisher's method, all but the second of these correlations are significant; the last is highly significant.)

It may be asked what possible meaning a factor can have that extends over such diverse objects as pictures, odours, colours and polygons. One might surmize that we are dealing here with a perceptual factor that derives directly from the nature of the nervous system; however, any positive answer to this question would involve too much speculation to be very helpful.

We have compared 'T', the general factor of aesthetic appreciation, with 'g', the general factor of cognitive ability. It may well be asked if these two factors are not at bottom one and the same thing, 'T' being merely the manifestation of 'g' when aesthetic material is used. In order to investigate this possibility, an intelligence test of the verbal type was given to the subjects,[1] and the correlation between 'T' and the scores in the intelligence test found. With a standard error of $\pm 0{\cdot}23$, the correlation was only 0·25, i.e. not significant statistically. It seems, therefore, that 'T' cannot be explained entirely in terms of general

[1] The test used has been described briefly by White(17).

cognitive ability, even though from the results of other workers it would appear that there is a slight positive correlation between intelligence and aesthetic appreciation.

Summarizing, we may state our main conclusions. First, evidence was found for a general, objective factor of visual aesthetic appreciation, which is independent of teaching, tradition, and other irrelevant associations. This factor is not identical with '*g*'. Secondly, this factor was found to extend over other modalities as well, correlating with a general factor derived from the ranking of odours. Thirdly, when this general factor was eliminated, a bipolar factor was found to emerge which seemed to divide the 'formal' from the 'representative' type of picture; a dichotomy lending support to theories long held by aestheticians.

REFERENCES

(1) DEWAR, H. (1938). *Brit. J. Educ. Psychol.* VIII, 29–49.
(2) WILLIAMS, E., WINTER, L. & WOODS, J. M. (1938). *Brit. J. Educ. Psychol.* VIII, 265–84.
(3) BURT, C. (1933). *The Psychology of Art,* ap. Burt, C., Jones, E. & Moodie, W. *How the Mind Works.* London: Allen and Unwin.
(4) DEWAR, H. (1937). Thesis in the Library of the University of London.
(5) MOORE, H. T. (1917). *Pain and Pleasure.* New York.
(6) EYSENCK, H. J. (1939). *J. Exp. Psychol.* XXV, 650–4.
(7) GARRETT, H. E. (1933). *Statistics in Psychology and Education.* New York: Longmans and Co.
(8) THURSTONE, L. L. (1935). *Vectors of Mind.* University of Chicago Press.
(9) THURSTONE, L. L. (1938). *Primary Mental Traits.* University of Chicago Press.
(10) EYSENCK, H. J. (1939). *Brit. J. Educ. Psychol.* IX, 270–5.
(11) HARTOG, P., RHODES, E. C. & BURT, C. (1936). *Marks of Examiners.* London Macmillan and Co.
(12) WELLS, F. L. (1908). In *Essays in Honor of William James.* New York: Longmans and Co.
(13) KANT, I. (1792). *Critik der Urtheilskraft.* Transl. J. H. Bernard (1914). London: Macmillan and Co.
(14) PORENA, M. (1905). *Che cos è il bello?* Milano.
(15) FRY, R. (1937). *Vision and Design.* Penguin Books.
(16) BELL, C. (1914). *Art.* London: Chatto and Windus.
(17) WHITE, H. D. J. (1931). *Brit. J. Educ. Psychol.* I, 279–95.