The aesthetics of simple figures

I. C. McManus

After reviewing the literature on experimental rectangle aesthetics ('the golden section') it was concluded that all of the effects demonstrated in the literature were dubious, either due to methodological limitations or inadvertent experimenter bias, these defects being compounded by most studies only considering population preferences and ignoring individual differences in preference. A series of experiments is described in which highly significant and temporally stable, but somewhat idiosyncratic, individual preferences were found. A taxonomy of these preferences, as well as of those for three separate types of triangle preference, is provided, based on factorial analysis. Two clear factors were demonstrated, one based on the square, and the other upon a proportion similar to that of the golden section.

> SOCRATES The beauty of figures which I am now trying to indicate is not what most people would understand as such, not the beauty of a living creature or a picture; what I mean...is something straight or round, and the surfaces and solids which a lathe, or a carpenter's rule and square produces from the straight and round...Things like that, I maintain, are beautiful not, like most things, in a relative sense; they are always beautiful in their very nature, and they carry pleasures peculiar to themselves... *Philebus*, 51.c

The problem of proportion has seriously concerned aesthetics for two and a half millennia. At its simplest the problem reduces to the question, What is the most harmonious manner in which to arrange a small number of lines or other modular elements? Two major theoretical approaches can be found (further details of which may be found in the reviews by Arnheim, 1955; Wittkower, 1960; Panofsky, 1970; Zusne, 1970; Berlyne, 1971).

The older approach emphasizes the integers and their relations. It is shown *par exellence* in the Pythagorean analysis of musical intervals, where the most harmonious pairs were found to be in the proportions 1:2, 2:3, etc.... . This system was codified by Vitruvius in his *De Architectura* of circa 30 B.C. The early Renaissance used the same aesthetic, it being clearly shown in Leone Battista Alberti's *Ten Books on Architecture*, published in 1485. This point is the historical watershed between two aesthetics; 24 years later Luca Pacioli in his *De Divina Proportione*, (1509) proposed that the fundamental proportion in aesthetics was the golden section. The fascinating mathematic properties of this geometric figure have been well described by Schooling (1914), Archibald (1920) and Huntley (1970). Pacioli's theory spread rapidly, influencing Dürer (Brion, 1960), Ramus, and Kepler (Sarton, 1951).

Classical and Renaissance aesthetic theorizing had been essentially *a priori* and prescriptive in its approach to aesthetics. Plato's *Timaeus* had suggested that within the properties of numbers themselves could be found the essence of the universe, and aesthetics was seen as a branch of this numerical cosmogony. This attitude was modified after the Renaissance, so that the origin of the numbers themselves was partly empirical. Thus Alberti claimed to have measured the actual proportions of human bodies, and Dürer also carried out much research on human proportion. Even Burke (1757) accepted that such natural proportion might compel him to accept a system of numerical aesthetics. As a part of the 19th century German revival in aesthetics, Zeising (1854, 1855), and also Henzlmann (1860) suggested that because the golden section and other ratios could be found empirically within nature then the ratio must (and the argument is still prescriptive) be of aesthetic significance.

Gustav Fechner, in his *Vorschule der Aesthetik* (1876), is usually represented as attempting to remove the prescriptive part of the argument and to actually measure aesthetic preferences themselves. There are however reasons for believing that Fechner also had a strong prescriptive wish which modified his experiments.

As a part of his new aesthetics 'from below' instead of in the traditional manner, 'from above', Fechner proposed three experimental methods. The most widely used is the 'method of choice', in which subjects are allowed to choose between several alternatives, selecting that which they feel is the most beautiful, or the most pleasant. In his most famous experiment Fechner seated subjects in front of a series of 10 rectangles of various width–length ratios and simply asked them to choose the one which they liked best and the one which they liked the least (Fechner, 1876). The modal preference was indubitably for the golden section, although the spread around the section was wide and there was a hint, particularly in Lalo's (1908) replication of the experiment, of a secondary peak at the square. The experiment has been enormously influential, being accepted by many non-scientists, and indeed by many psychologists, as incontrovertible scientific proof of the superiority of the golden section. It is therefore worth looking further at Fechner, and his motivations for studying this particular subject, as well as at some of his related research.

From an early age Fechner had waged a long war against the growing materialism of the 19th century, and this is partly manifested in his works on life after death, and on the mental life of plants (Fechner, 1835, 1848). His psychophysical researches were inspired when 'lying in bed on the morning of the 22nd October, 1850, he saw the vision of a unified world of thought, spirit and matter, linked together by the mystery of numbers' (Brett, 1921). His fascination with numerical aesthetics had been revealed earlier when he had published his thoughts on the form of angels and had concluded that they must be spherical, for the sphere was the most perfect of forms (Fechner, 1825; Boring, 1940). To such a man we may speculate that the mathematical properties of the golden section would represent a useful link between the harmony of nature and the world of the spirit. Whilst it is not possible to accuse Fechner of direct, nefarious alteration of his experimental results so that his data fitted with his prior theories, we may speculate as to how much Fechner, consciously or subconsciously, produced experimental circumstances which would tend to give him his desired results. Godkewitsch (1974) and Piehl (1976) have shown that the method of rank ordering is very sensitive indeed to artifacts, both of experimenter expectancy and subject expectation, as well as to the range of stimuli presented, both midpoint and anchoring tendencies being found (although Benjafield, 1976, suggests that some of Godkewitsch's own results may themselves be artifactual). Fechner's subjects were not selected at random, and it is quite feasible, particularly given his rejection elsewhere of double-blind methods (Fechner, 1860; David, 1968, pp. 16–17), that his 'cultured' subjects were well aware of the intentions of the experimenter. As Godkewitsch put it: 'In Fechner's study the subjects, asked to choose the most pleasing rectangle, often waited and wavered, rejecting one rectangle after another. Meanwhile the experimenter would explain that they should carefully pick a rectangle whose ratio between its sides could on the average be considered as most satisfying, harmonic and elegant'.

In summary we cannot accept Fechner's experiment as adequate proof of a general population preference for the golden section. It is also noteworthy that Fechner carried out experiments on the aesthetics of ellipses and, having failed to find a preference for the golden section, did not publish his results, these being found posthumously in his unpublished papers (Witmer, 1894).

As far as Fechner's other methods are concerned, he did not apparently carry out any experiments using the 'method of production', that is, asking subjects to draw or construct rectangles of the most pleasing proportions.

Fechner's 'method of application' was to consider works of art or other artifacts, and to examine the proportions used in their construction. He himself found that the mean height–width ratio of paintings was removed from the golden section, although he did not give actual distributions, only means (Fechner, 1876).

Fechner's principal experiment has been replicated several times. Studies by Lalo (1908), Thorndike (1917), Thompson (1946), Shipley *et al.* (1947), Nienstedt & Ross (1951), Eysenck & Tunstall (1968) and Berlyne (1971) have all used variants of the 'method of choice', simultaneously presenting a series of rectangles to a subject and asking him to rank them in order of preference, and have obtained broadly similar results to those of Fechner. However the method still has methodological defects (Godkewitsch, 1974; Piehl, 1976) and it must therefore be concluded that despite the replications the method is itself inadequate for the analysis of the problem of rectangle aesthetics.

Haines & Davies (1904) asked subjects to look at a single stimulus at a time and to 'accept or reject it'. They found large variations both within and between their small number of subjects. Hintz & Nelson (1970) used a method of 'successive approximation': this very dubious technique, which they also used in their 1971 study, would appear to be open to severe methodological criticism, not the least of which is that it makes no provision for subjects to have more than one preference. Amongst the 'methods of choice', only that of Weber (1930) seems to be devoid of serious methodological criticisms, since he used the method of paired comparisions, whereby a judgement is made separately for each pair of stimuli. Since there is good evidence that ranking tends to be a process of successive (but limited) paired comparisons (Russo & Rosen, 1975), this would seem to represent a solution to the problem of method, although even here we cannot be sure that range or anchor effects are not of significance.

Piehl (1978) also used the method of paired comparison. However he used only seven stimuli (and hence only 21 pairs) and his results are difficult to interpret due to the massive size of the standard deviation relative to the differences between means. Piehl also fails to state whether his stimuli were 'horizontal' or 'vertical'.

Whilst all of the studies of rectangle preferences are open to objection it is perhaps worth pointing out that in some cases correlations with other factors have been found, which may not be entirely attributable to methodological artifact. Both Weber (1930) and Eysenck & Tunstall (1968) found a tendency on repeat testing for longer, thinner rectangles to be preferred. In addition Eysenck & Tunstall found a slight tendency for introverts to prefer longer, thinner figures. Young children (mean age 3·7 years) have no consistent group preferences (Thompson, 1946), and as they grow older their preferences grow increasingly like that of adults. In the old (61–91 years) there is a tendency to prefer squarer figures (Nienstedt & Ross, 1951). Preferences tend to be shown more clearly if figures of constant area rather than constant side length are used as stimuli (Shipley *et al.*, 1947). There is a small correlation between the shape of the visual field and the preferred rectangle shape (Hintz & Nelson, 1970). Certain groups tend to show preferences for squares: Berlyne (1970) found this tendency amongst most Japanese subjects, but only *some* Canadian subjects, and Hintz & Nelson (1971) suggest that haptic preference in congenitally blind subjects is for squares (as opposed to a haptic preference for golden sections in sighted subjects).

Apart from rectangles there have been very few studies of other figures. Fechner himself looked at ellipses, and his results have been discussed earlier. Thorndike (1917) looked at triangles as well as crosses, and Lalo (1908) looked at crosses, and also dotted i figures (as had Fechner, see Witmer, 1894). Both sets of results are probably invalidated by the method of rank ordering. A notable exception to these methodological criticisms is the work of Austin & Sleight (1951) who looked at preferences for a series of isosceles triangles

by the method of paired comparisons and found a consistent population preference, although they also noted large individual differences and suggested that their preference curve primarily represented a curve of 'least dislike' (see Fig. 8).

The only other work of relevance is that studying regular polygons (e.g. Eysenck, 1968; Eysenck & Castle, 1970). They used as stimuli the figures of Birkhoff (1933) and asked subjects to rate each of them on a seven-point scale of 'Aesthetic pleasingness'. They factor analysed their results and found one factor of particular relevance to the present study, since it contained the square, a rectangle of ratio approximately 2:1, an equilateral triangle, an isosceles triangle, and a right-angled equilateral triangle.

From this survey of the experimental literature we are forced to conclude that there is really very little adequate evidence for any meaningful consistent population preference for simple figures, the only possible exceptions being the work of Weber (1931) and Austin & Sleight (1951). This is not however to say that there are no such preferences, although Godkewitsch (1974) and Piehl (1976) have suggested this (and Piehl, 1978, has since reversed his earlier decision). One major objection to all of the earlier work is that no study has been made of individual as opposed to population preferences. Population preferences often conceal large underlying individual differences. Those few studies where the authors have stressed the importance of looking at individuals (e.g. Haines & Davies, 1904; Thorndike, 1917) have been constrained by the lack of any adequate statistical test which will allow them to make meaningful statements about the preference of a single subject. No such test is possible with the limited data of a rank-ordering technique, but it is possible with the method of paired comparisons. A further problem is that no one has looked at preferences for rectangles, and also for other simple figures, such as triangles, *in the same subjects*. Similarly, with the limited exception of Weber (1931), there has been no attempt to make a reasonably long-term follow-up of individuals to find out how stable their preferences are. The present study attempts to remedy some of these defects.

A note on the description of stimuli

A rectangle may be described in terms of the ratio of the horizontal to the vertical side (if the area is constant). The difficulty with using the simple ratio, horizontal length (width, H) divided by vertical length (height, V), is that rotation through 90 degrees produces a figure whose shape or form is the same, but whose ratio is now the reciprocal of H/V. A more serious difficulty is that figures of equal intervals on this ratio scale are not perceptually equidistant (thus the difference between rectangles of ratio 1·1 and 1·2, is not psychologically the same as that between rectangles of ratio 0·2 and 0·3). In this report therefore the logarithm to the base 10 of the ratio (H/V) will be used throughout. This has the advantage that if a figure is simply rotated through 90 degrees then one merely has to alter the sign of the log. ratio (e.g. on rotation a rectangle of log. ratio 0·30 becomes one of log. ratio −0·30): the intervals are also approximately perceptually equidistant. The description of the ratio of isosceles triangles (of two types A and B, see Fig. 1), and right-angled triangles (triangle type C) will be in terms of the enclosing rectangle. Thus a golden section triangle is defined for these purposes as that drawn within a golden section rectangle (clearly this decision is in some sense arbitrary). The act of rotating a triangle of type A through 90 degrees is to make it of series B, and to also alter the sign.

The log. ratio of the golden section, (ϕ) 1·6180..., is 0·2089 and ratios of 2, 3 and 4 are respectively log. ratios 0·3010, 0·4771, and 0·6020.

Method

Altogether, three separate experiments have been carried out (Expts 1, 2 and 3). All used similar methods but differed in the details of presentation and of the particular stimulus types and values. In

each case a complete paired-comparison design was used. Subjects were shown, for a particular stimulus type, all possible pairs of stimuli and were asked to make a single response for each pair, stating their relative preference on a six-point rating scale (strong, medium, weak preference for stimulus A, weak, medium, strong preference for stimulus B). Minimal instructions were given to the subjects, they simply being asked to record which stimulus they 'preferred, or thought looked best'. No subject found this a strange task and they all settled down readily to the experiment proper after a brief practice run with about 10 pairs of stimuli. Subjects were asked to try as far as was possible to use all of the response categories. Most of the subjects were undergraduates at the University of Cambridge, none of whom was specializing in Fine Arts or Architecture, and very few of whom had ever heard of the golden section, or of its importance to aesthetics. Subjects were self-paced during the experiments, and they were encouraged to use immediate rather than considered impressions as judgements. Most subjects looked at each pair of stimuli for from 5 to 15 seconds before making a decision. In Expt 1 subjects were tested individually and in Expts 2 and 3 in pairs. Experimental sessions usually lasted from $\frac{3}{4}$ of an hour to $1\frac{1}{2}$ hours.

In Expt 1, 23 subjects were shown a series of 'horizontal' rectangles (i.e. rectangles of log. ratio > 0). Fifteen of these subjects also saw, usually in the same experimental session, but in a few cases on a separate occasion, a series of 'vertical' rectangles (i.e. of log. ratio < 0). A few subjects also saw, in separate experimental sessions, a few other series of stimuli (see Fig. 4). In Expt 2, 27 subjects were shown a series of 'mixed' rectangles (i.e. vertical and horizontal rectangles in the same series), then a series of upright isosceles triangles (triangles A, see Fig. 1), and then a series of isosceles triangles turned on their sides (triangles B, see Fig. 1). All subjects saw all stimulus types. Triangles A represented a replication of the experiment of Austin & Sleight (1951), except that a rating scale was used, and the stimuli were of slightly different proportions.
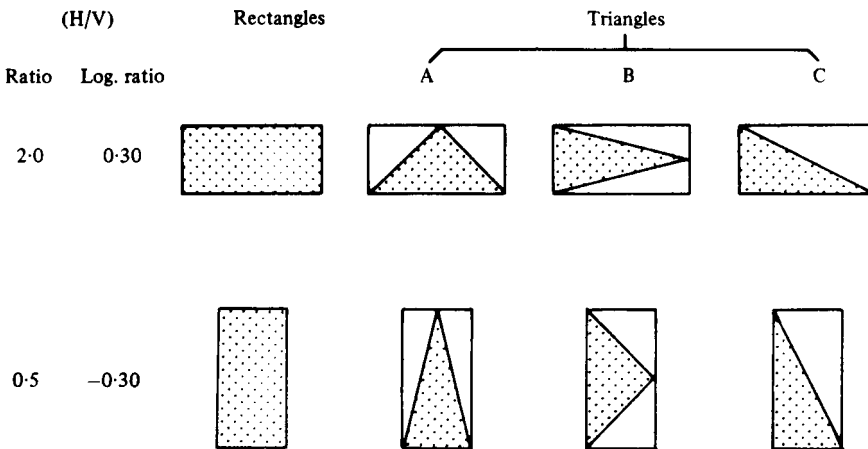

Figure 1. Definitions of ratios used for describing rectangles and triangles: see Results section for further description.

In Expt 3, 40 subjects saw a mixed series of rectangles, followed by a series of triangles A, and then triangles B, and finally triangles C (right-angled triangles – see Fig. 1). All subjects saw all four series of stimuli and, as in Expt 2, saw all stimulus types in the same order.

The particular stimulus values in each experiment and series, and the number of stimulus pairs shown, are given in Table 1.

The stimuli were shown by means of a pair of slide projectors and the members of each pair were shown side by side, the left–right positioning, as well as the overall order, being determined randomly, although the particular random order was the same for all subjects for any particular stimulus series. In Expt 1 the random order for horizontal rectangles was the same as that for vertical rectangles, the slides simply being rotated through 90 degrees in each projector. The stimuli consisted of solid white figures, all of equal area, projected against a dark background.

Table 1. The stimulus values used for each stimulus type in each experiment (see Fig. 1 for definitions of stimulus types). Values are expressed as the log. ratio × 100

Experiment	Stimulus type	No. of pairs	Stimulus values
1	Horizontal rectangles	105	0, 3, 6, 9. 12, 15, 18, 21, 24, 27, 30, 37·5, 45, 52·5, 60
1	Vertical rectangles	105	0, −3, −6, −9, −12, −15, −18, −21, −24, −27, −30, −37·5, −45, −52·5, −60
2	Rectangles	105	−52·5, −37·5, −27, −21, −15, −9, −3, 0, 6. 12, 18, 24, 30, 45, 60
3	Rectangles	105	−60, −45, −30, −24, −18, −12, −6, 0, 6, 12, 18, 24, 30, 45, 60
2	Triangles A	66	−47, −44, −40, −35, −30, −24, −18, −10, 0, 12·5, 30, 60
3	Triangles A	45	−60, −45, −30, −22·5, −15, −7·5, 0, 15, 30, 60
2	Triangles B	66	−60, −30, −12·5, 0, 10, 18, 24, 30, 35, 40, 44, 47
3	Triangles B	45	−60, −30, −15, 0, 7·5, 15, 22·5, 30, 45, 60
3	Triangles C	55	−45, −30, −22·5, −15, −7·5, 0, 7·5, 15, 22·5, 30, 45

A small number of subjects took part in the experiment on several occasions and their results have been included in each separate experiment in which they took part. Five of the subjects of Expt 1 took part in Expt 2, and four of these same subjects also took part in Expt 3; one subject in Expt 2, who had not taken part in Expt 1, also took part in Expt 3.

Analysis of results

Rectangles

Data were analysed by giving 5 points for a strong preference for a stimulus, 4 for a medium and 3 for a weak preference, and 2, 1, and 0 for a weak, medium or strong dislike respectively. Each single pair-comparison judgement therefore gave two numbers, a relative like for one stimulus, and a relative dislike for the other stimulus. These values were entered into an $n \times n$ matrix (the portion below the leading diagonal being the complement of the portion above). The leading diagonal was filled with zeros. Relative preferences for each stimulus were computed by taking the edge totals and then standardizing them by compressing them so that the maximum possible score was $+1·0$, and the minimum possible score was $-1·0$. This process was carried out both for individual subjects and also for groups of subjects, individual preference matrices having their respective cells summated.

Figure 2 shows the group preference for rectangles, the four different sets of stimulus values being plotted separately. Note that the ordinate is relatively expanded with respect to its possible range. Preference values may be compared within series, although not between series or between experiments. Several features are apparent in all of the curves. There is a dislike for the extreme ends of the spectrum, although this is stimulus-dependent and not range-dependent (compare Expt 1 with Expts 2 and 3). There is a preference for values around the golden section (ϕ) although in several cases the exact maxima are fairly discrepant from ϕ itself. The curves are also symmetric about log. ratio 0. In all
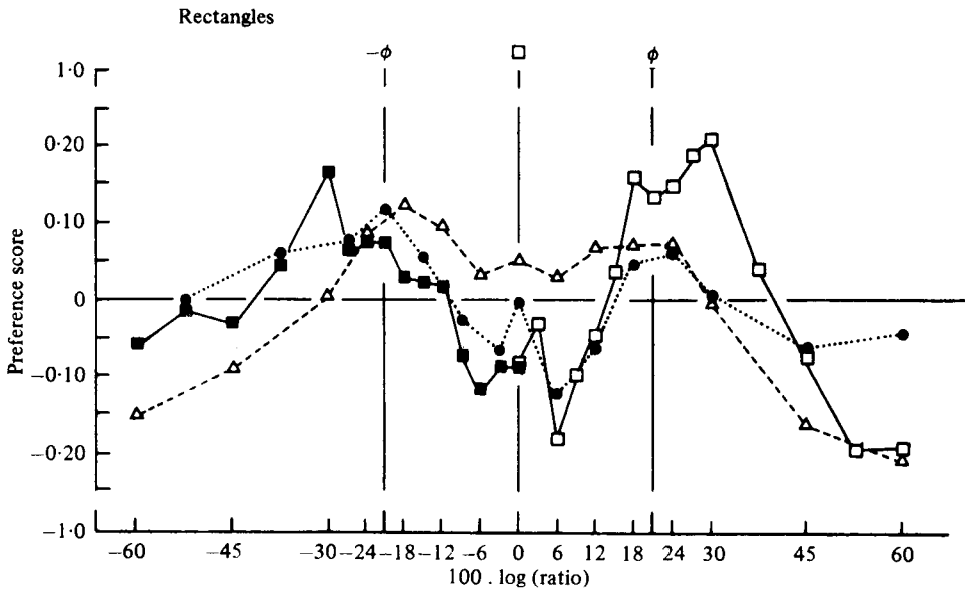
Figure 2. Population preferences for rectangles, results being plotted separately for the two parts of Expt 1, and for Expts 2 and 3. The vertical lines through the figure represent the positions of the square and the golden section. □–□, Expt 1H; ■–■, Expt 1V; ●··●, Expt 2; △---△, Expt 3.

experiments there is some evidence for a preference at about log. ratios of 0, i.e. the square, and in Expt 1 in particular this preference seems to be slightly shifted towards the right, a feature which might indicate the existence of the horizontal–vertical illusion, which has been shown to occur in solid rectangles (McManus, 1978). There is also perhaps a hint that the preferences around the golden section, both $+\phi$ and $-\phi$, are also shifted slightly to the right as well, although this conclusion is far from certain. The final point to note is the relatively small size of the population preference functions, the range being $+0.2$ to -0.2, in a possible range of $+1.0$ to -1.0.

In view of the relatively small effects it is desirable to have some form of statistical analysis against a null hypothesis. The null hypothesis is, of course, that the individuals in the experiment are simply responding at chance levels. To carry out a statistical analysis the individual preference matrices were converted into binary matrices, 0, 1, and 2 being collapsed into a value of 0, and 3, 4, and 5 being collapsed to a value of 1. These data were then analysed, either individually or as a group, by the methods described by David (1968). Two methods may be used. In the simplest (David, 1968, p. 38) the edge totals of the binary matrix may be tested for homogeneity. This has two limitations: the edge totals might be homogenous despite a significant micro-structure within the data matrix itself, and also the analysis takes no account of any possible trends along the edge, the particular order of the stimulus values not being taken into account. The second method (David, 1968, p. 25) based on that of Kendall & Babington Smith (1940) is more sensitive, taking account of the micro-structure of the data matrix itself. Consider three stimuli, P, Q, and R. Let a subject prefer (*p*)P to Q, and Q to R (i.e. P *p* Q, and Q *p* R). If he has consistent preferences we might expect that also P *p* R, whilst if he is merely responding at chance levels there should be an equal likelihood of R *p* P. Triads of the form P *p* Q, Q *p* R, P *p* R may be described as logical, transitive or consistent triads, whilst those of the form P *p* Q,

Q *p* R, R *p* P may be described as illogical, intransitive or circular triads. The number of circular triads within a data set is a sensitive index of the degree of consistency of the responses. The method does not however take note of possible trends in the data due to ordering of the stimulus values, and thus is still an inherently conservative test. The application of a test which takes into account the ordered nature of the stimuli, such as that of Jonckheere (1954), is also unsatisfactory since there are no *a priori* orders which are intuitively reasonable. Hence the data may be only tested for trend *a posteriori* and this is statistically unsatisfactory.

Using the method of circular triads for the combined horizontal rectangle preferences of Expt 1, the population value of U is 0·073 (possible range $= 1·0$ to $-0·043$) (the maximum value of U is of course 1, when all subjects agree completely, whilst the minimum value cannot be $-1·0$, for complete disagreement between subjects is not logically possible, and the particular minimum must be calculated for the particular preference matrix). For the horizontal rectangles the value of U is significantly different from chance ($\chi^2 = 304·5$, d.f. $= 120$, $P < 0·001$). For the vertical rectangles of Expt 1 there is no significant degree of inter-subject agreement by this test ($U = -0·024$, range of $U = 1·0$ to $-0·067$, $\chi^2 = 90·1$, d.f. $= 130$, n.s.). The results for Expt 2 are just significantly different from chance ($U = 0·009$, possible range of $U = 1·0$ to $-0·036$, $\chi^2 = 145·4$, d.f. $= 117$, $P < 0·05$), whilst those for Expt 3 are significant ($U = 0·049$, possible range of $U = 1·0$ to $-0·025$, $\chi^2 = 326·0$, d.f. $= 113$, $P < 0·001$). In summary it would seem that there probably are population preferences for rectangles, but that these effects are small.

A small population effect may be due either to an overall weak preference, or might be due to a strong preference within each subject, with these preferences being sufficiently different to cancel one another out when summed.

Figure 3 shows results for six individual subjects. All except subject 42 are significantly different from chance ($P < 0·001$) by the method of circular triads described above. It is clear from these individual preference functions that there is a wide range, and that the individual effects are of far greater magnitude than the population effect, preferences of close to $+1$ and -1 being reached in several subjects. Clearly the majority of the population preference function is a result of unjustified addition of qualitatively unlike individual functions.

The majority of individual preference functions are significantly different from chance. Table 2 shows the number of circular triads found in the three experiments. Clearly it is not possible for *all* triads to be circular. Kendall & Babington Smith (1940) demonstrated that for a 15×15 paired comparison matrix one can obtain a maximum of 140 circular triads (out of a total of 455 triads) and that chance alone would produce a modal value of 120 triads. A total of less than 96 triads is significant at the 5 per cent level, less than 81 at the 1 per cent level, and less than 72 at the 0·1 per cent level. From Table 2 it is clear that for all three experiments the majority of subjects are producing significantly non-random preference matrices.

The analysis of circular triads, as described above, has taken no account of the strength of the subjects' preferences, being based on the binary preference matrix. However we might reasonably expect that if subjects produce circular triads as a result of genuine errors, then when they do so they should make weaker judgements than when they are making non-circular triads. Let us give 1 point for a weak preference (in either direction), 2 for a medium preference, and 3 for a strong preference. Of course any single preference judgement is a part of a large number of triads, both circular and non-circular: nevertheless we may still calculate the average response strength in each of the triad types, remembering that such a test will not be maximally sensitive. If we consider, for each subject individually, the ratio of the strength in non-circular triads to that in circular triads
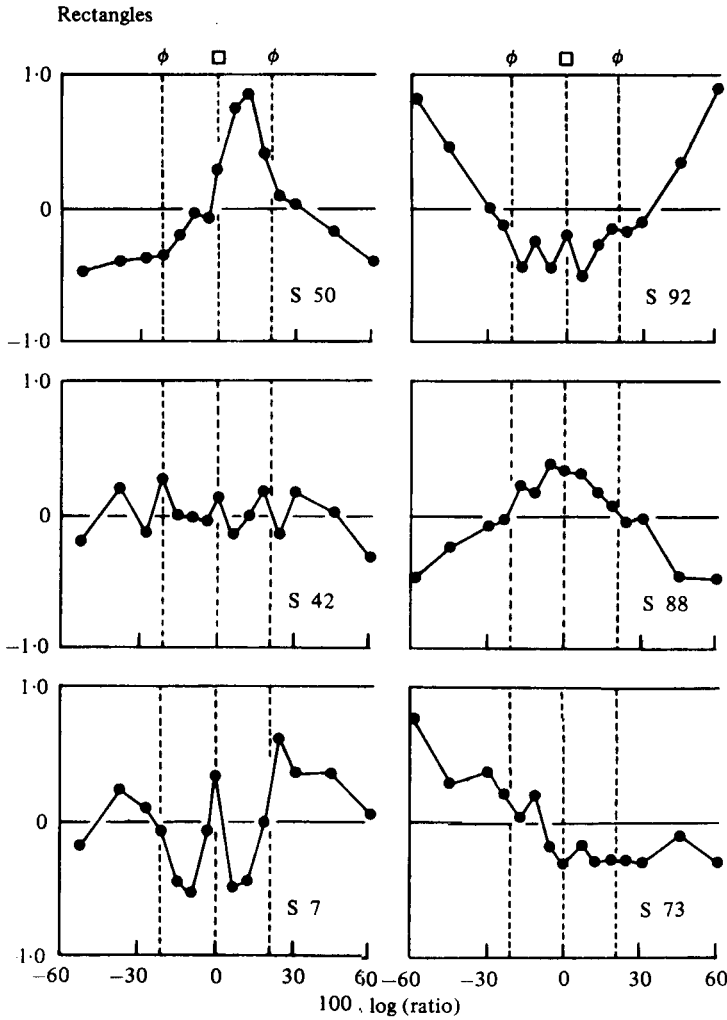
Figure 3. Individual rectangle preference functions for six subjects, subjects 7, 42 and 50 being from Expt 2, and subjects 73, 88 and 92 from Expt 3. All the preference functions, except that of subject 42, are significantly different from chance with $P < 0.001$. Subject 42's rectangle preferences are indistinguishable from chance: however she also carried out triangle preference experiments and produced highly significant results, a finding which occurred in several other subjects. The examples have been chosen for their range rather than in proportion to their actual rate of occurrence.

Table 2. The number of subjects producing various numbers of circular triads in the rectangle experiments

n triads sig:	$0 \rightarrow 71$ $P < 0.001$	$72 \rightarrow 80$ $0.001 < P < 0.01$	$81 \rightarrow 96$ $0.01 < P < 0.05$	$97 \rightarrow 140$ n.s.	n
Expt 1H	16	1	2	4	23
Expt 1V	12	0	2	1	15
Expt 2	19	0	3	5	27
Expt 3	30	1	6	3	40
Total	77	2	13	13	105

(x), then if the above hypothesis is correct then the value of x should be greater than 1. Overall of 104 subjects viewing rectangles, 82 had values of $x > 1$, and only 22 had values $\leqslant 1$ ($\chi^2 = 33.47$, d.f. $= 1$, $P < 0.001$). For an individual subject the significance of the difference between triad types may be found by comparing the frequencies of the three response categories in a 3×2 contingency table. Table 3 shows that overall 48 subjects (46·1 per cent) had significantly high values of x as compared with only 4 (3·8 per cent) subjects with significantly low values of x.

Table 3. Strength of preference judgements used in circular and non-circular triads, in rectangle experiments. Let $x = $ (non-circular score)/(circular score). One subject produced no circular triads at all and has been excluded from this analysis

Experiment	$x < 1$ $P < 0.05$	$x < 1$ n.s.	$x > 1$ n.s.	$x > 1$ $0.05 > P > 0.001$	$x > 1$ $P < 0.001$	n
1H	0	6	1	6	10	23
1V	2	2	5	4	1	14
2	1	5	12	3	6	27
3	1	5	16	10	8	40
Total	4	18	34	23	25	104

It thus seems that the preference functions of individual subjects are statistically highly significant. These functions are however only of real interest if they can be shown to be stable with respect to time. Experiments 1 and 3 of this study were carried out at an interval of $2\frac{1}{4}$ years. Four subjects took part in both experiments, and their preference functions are shown in Fig. 4. Note that these four subjects did not use, in this particular part of Expt 1, the stimulus values reported in Table 1, but a 'mixed' series, which provides greater compatibility with Expt 3. It is clear, given the range of individual preferences shown in Fig. 3, that these four preference functions show a reasonable temporal stability.

Triangles

The analysis of the triangle preference functions is essentially similar to that of the rectangles described in the previous section. Figure 5 shows the preference functions for each triangle type. All of the functions are significantly different from chance by the circular triad method (David, 1968, p. 25); see Table 4 for statistical analysis.

Examination of the preference functions of Fig. 5 shows several features of note. For type A triangles the results compare very closely with those of Austin & Sleight (1951); preference functions thus seem to be stable across 23 years and two continents. For all of the triangle types the preference function seems to be unimodal (unlike the case of the rectangles), but like the rectangle preference functions, the magnitude of the population preferences is small in comparison with their possible range of $+1$ to -1. For triangles of type A and B the golden section seems to be of some importance, but interestingly only at $-\phi$ for type A and $+\phi$ for type B: the implication is that it is the form of these triangles which is important, rather than the shape of their enclosing rectangle. The function for triangles C is fairly symmetric, but is so flat-topped that it is difficult to know exactly where the maximum is located, or even whether the curve is unimodal.

Analysis of the circular triads from individual subjects reveals that the majority of individual subjects have highly significant preference functions and that, as in rectangle
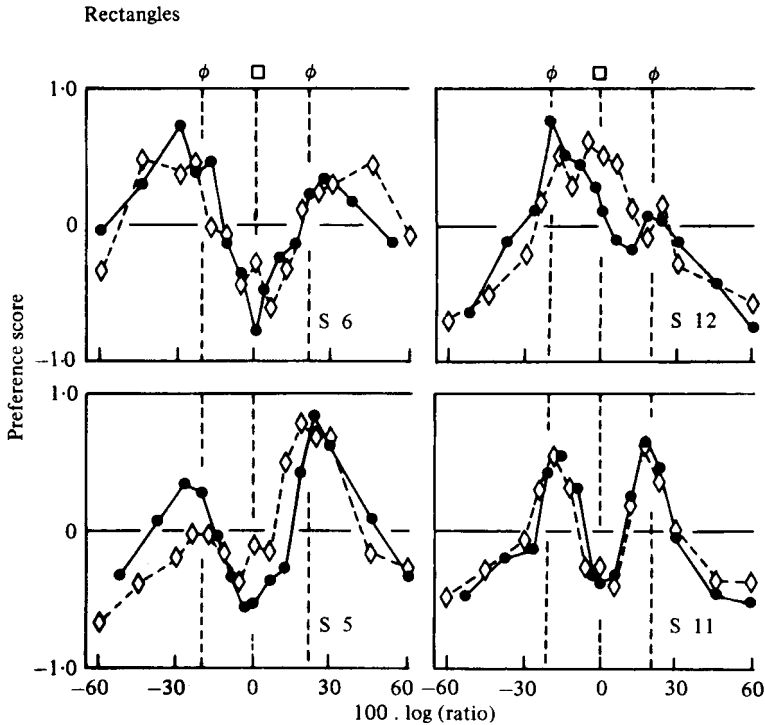
Rectangles

Figure 4. Test–retest results for rectangle preferences for four subjects. The preference functions indicated by the solid circles were measured first, whilst those with open diamonds were measured about 2 years later. The retest experiment used the rectangles already described for Expt 3. The test results however were obtained at the same time as Expt 1 was carried out but used the stimuli described for Expt 2; as the experiment was still at a somewhat early stage at that time, subject 6 actually used the inverse of the Expt 2 rectangles (i.e. rotated 90 degrees). These four subjects also carried out Expt 1 proper and their results obtained with those stimuli are included in the results of Expt 1, as described elsewhere in this paper.

preferences, there is a wide range of individual difference in preference functions: considerations of space preclude the description of individual variation for all triangles elsewhere in this paper. ●——●, test; ◇---◇, retest.

Individual differences in preference functions

The small population preference for rectangles (Fig. 2) contrasts strongly with the far greater magnitude of individual preference functions (Figs 3, 4), and implies that inter-subject differences are larger than inter-subject similarities. Visual scrutiny of the individual preference functions did not suggest any obvious taxonomy for these variations, and a multivariate statistical technique was therefore used to identify the underlying structure.

Consider rectangles in Expt 3. Each subject made preference judgements on 105 pairs of stimuli. The preferences for each pair of subjects were correlated by a pairwise comparison of the judgements of each subject. These correlation coefficients were therefore independent of the ordered nature of the 15 stimulus values. The 40×40 correlation matrix thus produced was then factor analysed (by means of the FACTOR program of the SPSS statistical package, Nie *et al.*, 1975), the first eight factors being extracted, and then a
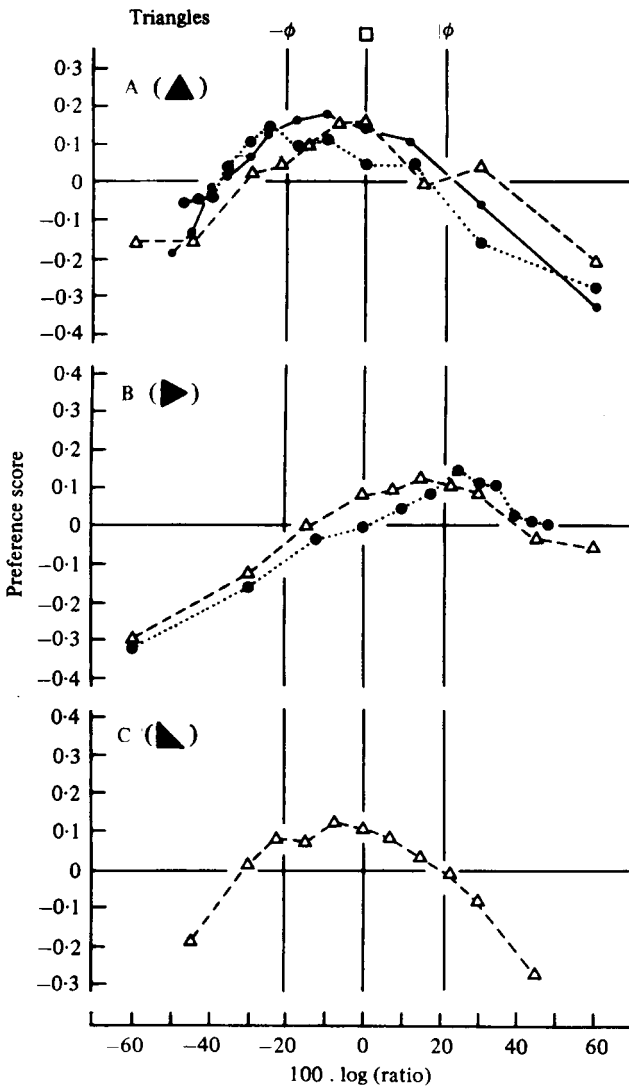
Figure 5. Population preference functions for Expts 2 and 3 for three separate types of triangles.
●···●, Expt 2; △---△, Expt 3; ●—●, Austin & Sleight (1951).

Table 4. Analysis of circular triads for population triangle preferences shown in Fig. 6:
method is that of David (1968, p. 25)

| Stimulus type | Experiment | U | Possible U range | | χ^2 | d.f. | P |
			Max	Min			
Triangles A	2	0·058	1	−0·038	181·9	74	< 0·001
Triangles A	3	0·077	1	−0·026	191·6	48	< 0·001
Triangles B	2	0·047	1	−0·038	161·7	74	< 0·001
Triangles B	3	0·065	1	−0·026	168·3	48	< 0·001
Triangles C	3	0·059	1	−0·025	192·4	59	< 0·001

varimax rotation used, producing orthogonal factors. To analyse the nature of these factors the factor loadings of each subject on a particular factor were multiplied by the subject's preference matrix, and the resultant weighted matrices then summed and the edge totals taken. These edge totals were then standardized so that the absolute total was equal to 2. These standardized edge totals were then plotted against rectangle shape, and the results inspected. A similar process was used for the other stimulus types in each of the experiments.

It is clearly of interest to know whether when a subject chooses a particular type of rectangle he will also choose a particular type of triangle A, triangle B, etc. To examine this the factor loadings of each subject of each of the eight factors of the four stimulus types in Expt 3 were intercorrelated across subjects, so that the resultant 32×32 matrix contained the intercorrelations between factors, e.g. between the third factor on triangles A and the fourth factor on triangles B, etc. This correlation matrix was then factor analysed
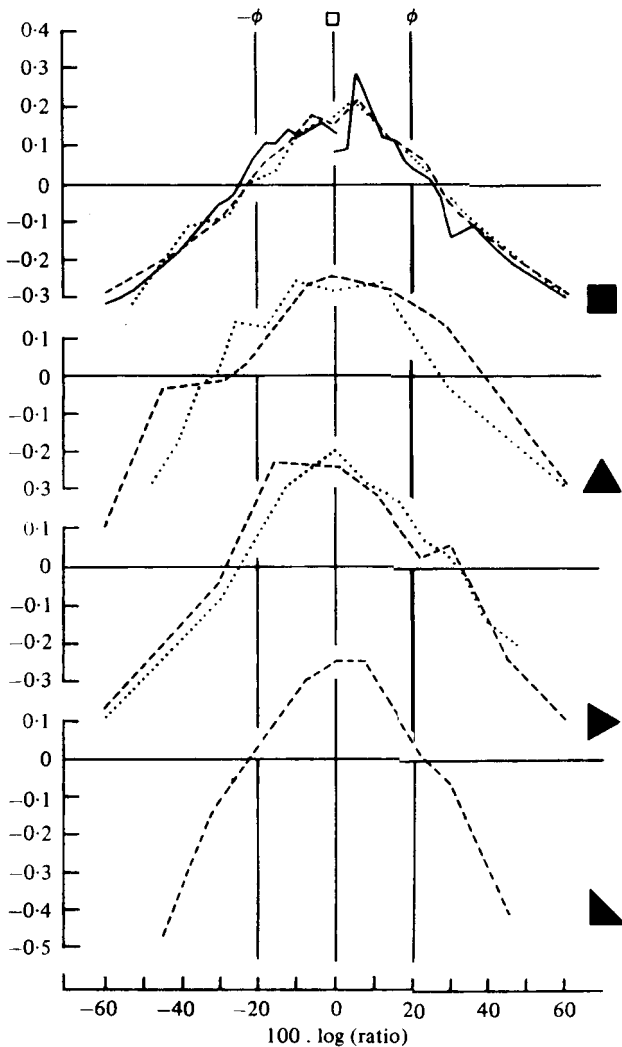

Figure 6. Factor alpha for the four stimulus types, for all experiments. The ordinate is arbitrary, being constructed such that the total absolute deviation of all the points from the abscissa zero should be $+2$. —, Expt 1; $\cdots$, Expt 2; ———, Expt 3.
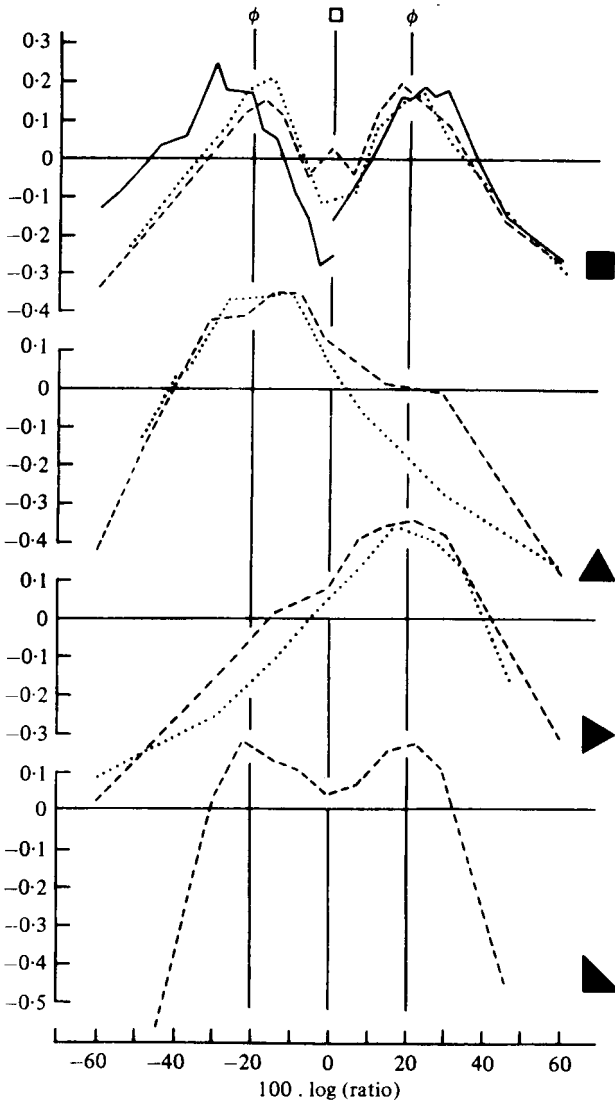
Figure 7. Factor beta: otherwise as for Fig. 6. —, Expt 1; · · ·, Expt 2; – – –, Expt 3.

using the same program, and the first eight factors extracted, and then rotated by a varimax rotation. From this analysis it became readily apparent which of the main factors of the individual stimulus type factor analyses were interrelated. Whilst these relationships were usually clear the process was not always unambiguous, particularly when the process was repeated for Expt 2, with its more limited numbers of subjects.

To clarify the interrelationship between stimulus types the individual subjects' data from Expt 3 were again intercorrelated but this time not just for one stimulus type at a time, but for all four stimulus types. Each correlation was therefore based upon 250 pairs of judgements. This 40 × 40 inter-subject correlation matrix was factor analysed and rotated to varimax orthogonality. From the loadings of each subject on the orthogonal factors the underlying preference functions were determined for each stimulus type separately. A similar process was carried out for Expt 2. The eigen values for the first 10 factors from the analysis of Expt 3 are 11·48, 4·08, 3·08, 1·81, 1·49, 1·30, 1·13, 1·09, 0·97, 0·94. A
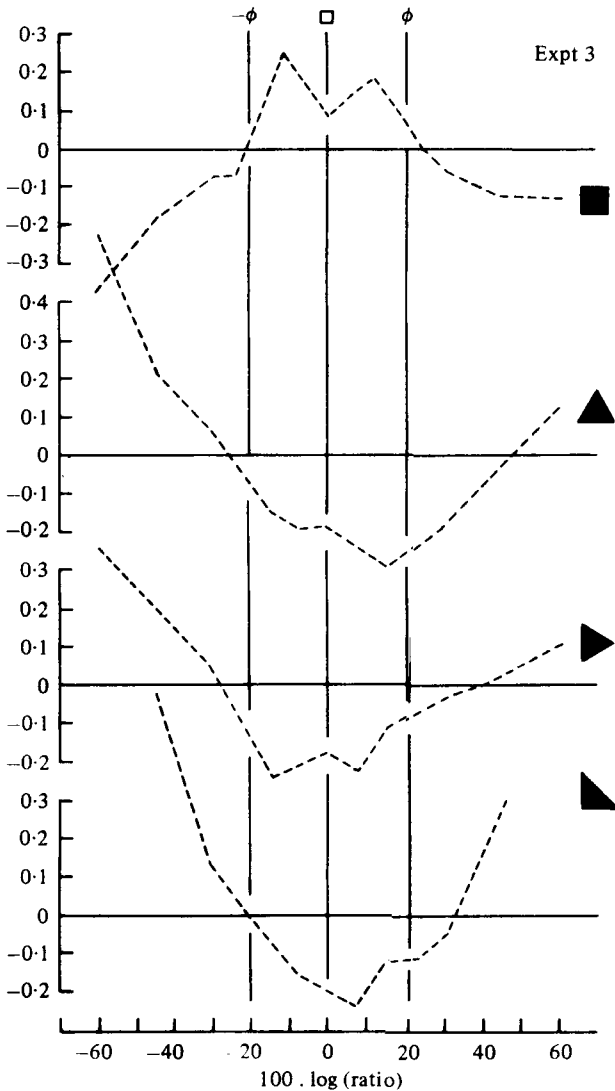
Figure 8. Factor gamma for Expt 3 only: otherwise as for Fig. 6.

'Scree-slope analysis' (Child, 1970) suggests that the first two factors are highly significant and the next two possibly so. The rest of the factors are probably too small, even if real, to be of any interest. It is possible of course that with a larger sample further factors would appear. The identification of similar factors in the three experiments was carried out by inspection. The first two factors of Expt 3 are readily identifiable in Expts 1 and 2, but the identification of the third and fourth factors is not clear and these have been omitted from this paper.

The first factor identified in Expt 3 has been called factor alpha, and is shown in Fig. 6. In its positive form it is a preference for squares and their triangular derivations, and a dislike for all other figures. In Expt 1H there is a suggestion of a horizontal–vertical illusion, although its magnitude is rather large. The triangle preferences are all broadly identical with each other.

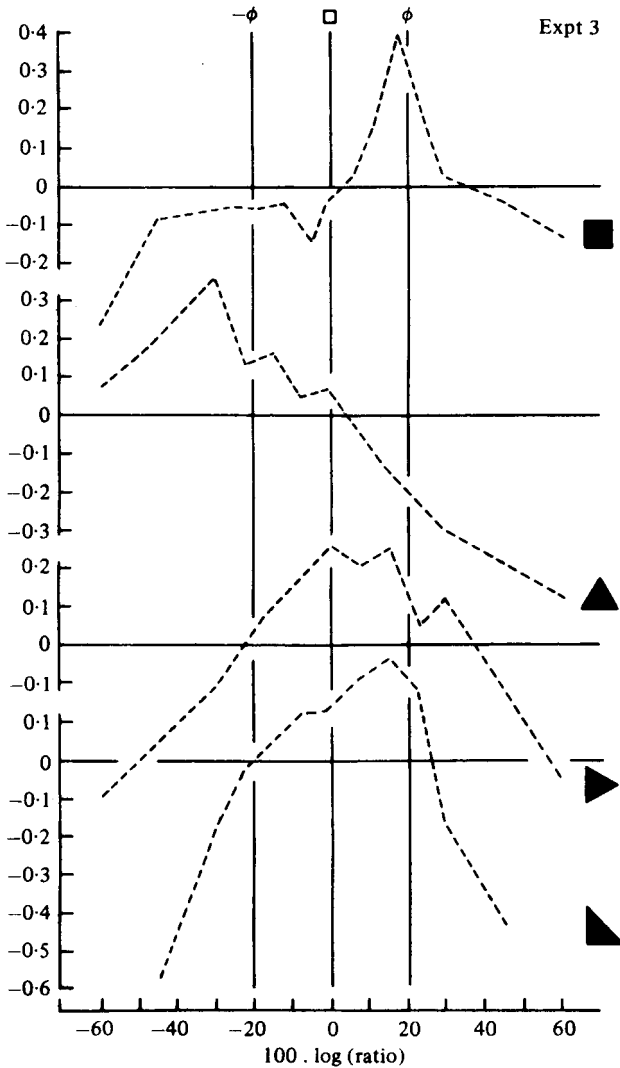Factor beta is the second factor isolated from Expt 3, and is shown in Fig. 7. It is rather

Figure 9. Factor delta for Expt 3 only: otherwise as for Fig. 6.

more interesting than factor alpha, being bimodal in the case of rectangles and right-angled triangles, and also showing a strong suggestion of being related to the golden section, both positive ($+\phi$) and negative ($-\phi$). Triangles A and B show similar curves, but differ from the other two stimulus types in being unimodal. They are also out of phase, triangles B being the mirror-rotation of triangles A around the 'square'. As noted earlier with the preference for triangles A and B (Fig. 5) the implication is that it is the form of the triangles which matters and not the orientation of the enclosing rectangle.

Table 5 shows the loading of each subject of Expt 3 on the first four factors. Twenty-two (55 per cent) of the subjects have a significant (> 0.3 or < -0.3) (Child, 1970) loading on the alpha factor, and 18 (45 per cent) have a significant loading on the beta factor. Only 11 subjects (27.5 per cent) have non-significant loadings on both factors. Eleven (27.5 per cent) subjects have significant loadings on both factors, the majority (8) having a positive

Table 5. Experiment 3: the number of individuals with various degrees of loading on the varimax rotated factors. Loadings in italics may be regarded as significant

	Loadings										
Factor	$-1{\cdot}0 \to$ $-0{\cdot}7$	$0{\cdot}69 \to$ $-0{\cdot}5$	$-0{\cdot}49 \to$ $-0{\cdot}30$	$-0{\cdot}29 \to$ $-0{\cdot}10$	$-0{\cdot}09 \to$ $+0{\cdot}09$	$0{\cdot}10 \to$ $0{\cdot}29$	$0{\cdot}30 \to$ $0{\cdot}49$	$0{\cdot}5 \to$ $0{\cdot}69$	$0{\cdot}7 \to$ $1{\cdot}0$	Total $< -0{\cdot}3$	Total $> 0{\cdot}3$
Alpha	0	4	3	5	6	7	3	8	4	*7*	*15*
Beta	0	1	2	4	11	7	7	4	4	*3*	*15*
Gamma	0	3	1	7	18	6	2	2	0	*4*	*4*
Delta	0	0	2	11	18	4	1	0	0	*2*	*1*

loading on both factors. It is important to note that seven subjects have *negative* loadings on factor alpha, and three subjects have negative loadings on factor beta: the preference functions for these subjects are therefore the inverse of those shown in Figs 6 and 7. It is these negative loadings which account for the relative flatness of the preference functions of Figs 2 and 5.

The last two factors (factor gamma and factor delta) are composed mainly of non-significant loadings with just a few subjects with significant loadings. It is difficult to interpret these factors with certainty. For interest and completeness they are shown in Figs 8 and 9, but their identification should be regarded as only tentative. Factor gamma is of interest in that it is almost identical to factor alpha except for the inversion of the triangle preference functions: the status of this finding is not at all clear, but it accounts for the partly ambiguous results obtained when the stimulus types were factored independently and then the loadings refactored, as described earlier. Factor delta is of particular interest for it is asymmetric around log. ratio zero, and approximates, in the case of the rectangle, to a unimodal golden section curve. Clearly it is necessary to have asymmetric preference functions of this type in order to account for the individual preference functions of the type shown by subjects 73, and possibly 50, in Fig. 3.

Conclusions

After an historical and experimental review it was concluded that the golden section phenomenon, particularly as delineated by Fechner, was probably unreliable and mainly artifactual. The paired-comparison technique is probably free from the artifacts of ranking methods; nevertheless with four different series of stimulus values, consistent preference functions were obtained for rectangles, these preferences being stable in several subjects over a period of $2\frac{1}{4}$ years. However population preferences were small in comparison with individual variation. After a moderate degree of statistical manipulation using multivariate analysis it was possible to produce an objective taxonomy of these individual variations and to produce at least two major factors which are readily interpretable and probably reliable, and also two other factors of probable significance and of some interest. It is presumed that the wide range of particular subject preference functions is a function of differential admixture of these several types of preference function. The other simple figures studied, three types of triangle, all bear a simple relation to the rectangle functions obtained.

It has been assumed throughout this study that the responses of subjects in this study truly represent 'aesthetic' responses: this may however be, at best, a tenuous assumption. The wide range of subject preference functions makes one speculate whether one is really dealing with some form of experimenter demand effect. Presumably since a subject has agreed to spend an hour doing the experiment he feels obliged to actually *do* something; this doing need not however represent aesthetic behaviour. Against this hypothesis are two items of evidence. Firstly the subjects claim that they *are* making aesthetic judgements, and that they feel the request to make such judgements is a reasonable one. Secondly, in those few subjects who have been retested over a 2-year period, there is evidence of a high degree of replicability (and also the subjects themselves claimed not to be able to remember their previous judgements).

It is perhaps reasonable therefore to assume that the results in this type of experiment represent some form of elemental aesthetic judgement of the type postulated by Socrates. If so it would be interesting to know why subjects differ, how constant their preference functions are over longer periods of time, and whether their particular preference functions correlate with other variables (shape of eye-field, personality variables, etc.). Also, does the *degree* of loading on particular factors vary within subjects, or correlate with other

variables. All of these items of information could give some clues as to the origin of the phenomenon, and its nature. The author has examined the data already presented in this paper in terms of the sex of the subject, and their particular area of study (arts vs. sciences) and has been able to find no significant links with the factors.

It would thus appear (admittedly somewhat to this author's surprise) that there is moderately good evidence for the phenomenon which Fechner championed, even though Fechner's own method for its demonstration is, at best, highly suspect owing to methodological artifacts. Whether the golden section *per se* is important, as opposed to similar ratios (e.g. 1·5, 1·6, or even 1·75), is very unclear, techniques at present not being accurate enough to make adequate measurements.

Acknowledgements

References

ALBERTI, L. B. (1485). *De re aedificatoria*. Florence. Translated by J. Leoni as, *The Ten Books of Architecture*. London: Tiranti, 1955.

ARCHIBALD, R. C. (1920). Notes on the Logarithmic Spiral, Golden Section, and the Fibonnaci series. In J. Hambidge (ed.), *Dynamic Symmetry: The Greek Vase*. Appendix, pp. 146–157. New Haven, Conn.: Yale University Press.

ARNHEIM, R. (1955). A review of proportion. *Journal of Aesthetics and Art Criticism*, **14**, 44–57.

AUSTIN, J. R. & SLEIGHT, R. B. (1951). Aesthetic preference for isosceles triangles. *Journal of Applied Psychology*, **35**, 430–431.

BENJAFIELD, J. (1976). The 'Golden Rectangle': Some new data. *American Journal of Psychology*, **89**, 737–743.

BENJAFIELD, D. E. (1970). The Golden Section and hedonic judgements of rectangles: A cross cultural study. *Sciences de l'Art*, **7**, 1–6.

BERLYNE, D. E. (1971). *Aesthetics and Psychobiology*. New York: Appleton-Century-Croft.

BIRKHOFF, G. (1933). *Aesthetic Measure*. Cambridge, Mass.: Harvard University Press.

BORING, E. G. (1940). *History of Experimental Psychology*, p. 279. New York: Appleton-Century-Croft.

BRETT, G. S. (1921). *A History of Psychology*, vol. III, p. 128. London: George Allen & Unwin.

BRION, M. (1960). *Albrecht Dürer: His Life and Work*. London: Thames & Hudson.

BURKE, E. (1757). *A Philosophical Enquiry into the Origin of our Ideas of the Sublime and Beautiful*. London: R. & J. Dodsley.

CHILD, D. (1970). *The Essentials of Factor Analysis*. London: Holt, Rinehart & Winston.

DAVID, H. A. (1968). *The Method of Paired Comparisons*. London: Griffin.

EYSENCK, H. J. (1968). An experimental study of aesthetic preferences for polygonal figures. *Journal of General Psychology*, **79**, 3–17.

EYSENCK, H. J. & CASTLE, M. (1970). Training in art as a factor in the determination of preference judgements for polygons. *British Journal of Psychology*, **61**, 65–81.

EYSENCK, H. J. & TUNSTALL, O. (1968). La Personnalité et l'esthétique des formes simples. *Sciences de l'Art* **5**, 3–9.

FECHNER, G. T. (1825). *Vergleichende Anatomie der Engel*. Leipzig.

FECHNER, G. T. (1835). *Das Buchlein vom Leben nach dem Tode*. Leipzig. Translated by H. Wernekke as, *On Life after Death*, Chicago, 1906.

FECHNER, G. T. (1848). *Nanna, oder über das Seelenleben der Pflanzen*. Leipzig: Leopold Bok. (See Fechner, 1946, for a partial translation.)

FECHNER, G. T. (1860). *Elemente der Psychophysik*. Leipzig: Breitkopf & Härtel. Translated by H. E. Adler as, *The Elements of Psychophysics*. New York: Holt, Rinehart & Winston, 1966.

FECHNER, G. T. (1876). *Vorschule der Aesthetik*. Leipzig: Breitkopf & Härtel.

FECHNER, G. T. (1946). *Religion of a Scientist*. London: Kegan Paul, Trench & Trubner.

GODKEWITSCH, M. (1974). The Golden Section: An artefact of stimulus range and measure of preference. *American Journal of Psychology*, **87**, 269–277.

HAINES, T. H. & DAVIES, A. E. (1904). The psychology of aesthetic reaction to rectangular forms. *Psychological Review*, **11**, 249–281.

HENZLMANN, E. (1860). *Théories des Proportions*. Paris: A. Bertrand.

HINTZ, J. M. & NELSON, T. M. (1970). Golden Section: Reassessment of the perimetric hypothesis. *American Journal of Psychology*, **83**, 126–129.

HINTZ, J. M. & NELSON, T. M. (1971). Haptic aesthetic value of the Golden Section. *British Journal of Psychology*, **62**, 217–223.

HUNTLEY, H. E. (1970). *The Divine Proportion*. New York: Dover Books.

JONCKHEERE, A. R. (1954). A distribution-free k-sample test against ordered alternatives. *Biometrika*, **41**, 133–145.

KENDALL, M. G. & BABINGTON SMITH, B. (1940). On the method of paired comparisons. *Biometrika*, 31, 324–345.

LALO, C. (1908). *L'Esthétique expérimentale et contemporaine*. Paris: Alcan.

MCMANUS, I. C. (1978). The horizontal–vertical illusion and the square. *British Journal of Psychology*, 69, 369–370.

NIE, N. H. et al. (1975). *Statistical Package for the Social Sciences*, 2nd ed. New York: McGraw-Hill.

NIENSTEDT, C. W. & ROSS, S. (1951). Preferences for rectangular proportions in college students and the aged. *Journal of Genetic Psychology*, 78, 153–158.

PACIOLI, L. (1509). *De Divina Proportione*. Venice: Paganinos de Paganinis de Brixia.

PANOFSKY, E. (1970). The history of the theory of human proportions as a reflection of the history of styles. In, *Meaning in the Visual Arts*. Harmondsworth, Middx: Penguin.

PIEHL, J. (1976). The 'Golden Section': An artefact of stimulus range and demand characteristics. *Perceptual and Motor Skills*, 43, 47–50.

PIEHL, J. (1978). The Golden Section: The 'true' ratio? *Perceptual and Motor Skills*, 46, 831–834.

RUSSO, J. E. & ROSEN, L. D. (1975). An eye-fixation analysis of multi-alternative choice. *Memory and Cognition*, 3, 267–276.

SARTON, G. (1951). *Isis*, 42, 47.

SCHOOLING, W. (1914). The ϕ progression. In T. A. Cook (ed.), *The Curves of Life*. Appendix, pp. 441–447. London: Constable.

SHIPLEY, W. C., BATTMAN, P. E. & STEELE, B. A. (1947). The influence of size upon preferences for rectangular proportions in children and adults. *Journal of Experimental Psychology*, 37, 333–336.

THOMPSON, G. G. (1946). The effect of chronological age on aesthetic preferences for rectangles of different proportions. *Journal of Experimental Psychology*, 36, 50–58.

THORNDIKE, E. L. (1917). Individual differences in judgements of the beauty of simple forms. *Psychological Review*, 24, 147–153.

VITRUVIUS POLLIO (1931). *On Architecture*. Translated by F. Granger. London: Loeb Classical Library.

WEBER, C. O. (1931). The aesthetics of rectangles and theories of affection. *Journal of Applied Psychology*, 15, 310–318.

WITMER, L. (1894). Zur experimentallen Aesthetik einfacher räumlicher Formverhaltnisse. *Philosophische Studien*, 9, 96–144.

WITTKOWER, R. (1960). The changing concept of proportion. *Daedalus*, 89, 199–215.

ZEISING, A. (1854). *Neue Lehre von den Proportionen des menschlichen Körpers*. Leipzig: Weigel.

ZEISING, A. (1855). *Aesthetische Forschungen*. Frankfurt am Main.

ZUSNE, L. (1970). *Visual Perception of Form*. New York: Academic Press.