EMPIRICAL STUDIES OF THE ARTS, Vol. 13(2) 131-148, 1995

FECHNER'S EXPERIMENTAL AESTHETICS AND THE GOLDEN SECTION HYPOTHESIS TODAY

HOLGER HÖGE
University of Oldenburg, Germany

ABSTRACT

Gustav Theodor Fechner started his "aesthetics from below" with an investigation of the golden section hypothesis because he was fascinated by the fact that a mathematic proportion should directly correspond to pleasingness. Thus, the golden section hypothesis fitted in his psychophysical approach, assuming that there is a correspondence between physical properties of stimuli and the sensations they cause. However, Berlyne believed that even minute variations in instruction or experimental arrangements may cause differences in the proportions preferred [1, p. 300]. Different from many investigations in the field the experimental manipulation to be reported here was to give different verbal criteria to our subjects. One of our experiments followed Fechner's method of production—i.e., subjects had to draw rectangles—and the other one was done using the method of choice—i.e., subjects had to sort rectangles. The results show that different criteria lead to different proportions in the material produced and sorted, respectively. Thus, preference judgements seem to be the outcome of a process of information processing by using both sources of information: the physical arrangement of the stimuli and the cognitively represented concept of the subject. However, under both conditions no preference for the golden section was found.

Fechner is not only called the father of experimental psychology, but he is also known as the founder of experimental aesthetics. As a real "Gelehrtenpersönlichkeit" (a word that is not so easy to translate: "an outstanding man of letters" might be a reasonable equivalent) he was a specialist in many fields and, thus,

doi: 10.2190/UHTQ-CFVD-CAU2-WY1C
http://baywood.com

tried to find unity in variety, because he developed his own way of uniting philosophy, medicine, physics, and finally psychology with that specific Fechnerian innovation: psychophysics.

Perhaps, Fechner had his first contact with aesthetics during his early scientific studies [2, p. 2]. when he read the works of Oken, a physiologist and philosopher of nature who believed that the senses are the "direct consensus of the nervous system with the world" ("unmittelbarer Consensus des Nervensystems mit der Welt," [3, pp. 106, 110]; cit. after [4, p. 241] and, of course, one may suppose that this was the first indication for Fechner that there is some correspondence between the outside and the inside world (cf. [5]).

Moreover, Fechner belonged to a private circle consisting of two theologists (Billroth and Grimmer), a mathematician (Müller), a publisher (Härtel), and a philosopher (Weisse); also, he was in contact with the poet Bettina von Arnim, and he was an active member of the Leipzig Art Association (Leipziger Kunstverein)—which may show his close connection with problems of aesthetics and art [6, p. v]. Of course, this is true for several researchers of those times but what makes the Fechnerian approach stand out is the fact that he gave the field of aesthetics a new direction—i.e., to start "from below" instead of "from above." These terms became widespread; aesthetics "from below" is now synonymously used with experimental aesthetics.

The aim of this approach was partly based upon the fact that Fechner was not content with the influence of philosophy on the natural sciences. Physics, he believed, was led astray by the philosophy of nature [6, p. 4]. He felt that in aesthetics philosophy was mainly following the rule of deduction—i.e., postulating what the terms desired—instead of looking at the reasons of how and why people had an aesthetic experience, an aesthetic judgement (cf. [7]). Therefore, he was convinced that systems of aesthetics are based on 'feet of clay' ("thönerne Füβe"; [6, p. 4]). With his new approach "from below" he started a line of thinking which has some similarity with the distinction between "bottom up" and "top down" processes.

Consequently, Fechner left his study to ask people about their judgements on works of art when the two versions of Holbein's Madonna were exhibited at the Dresden museum—this was the first time that an empirical method was used to investigate people's aesthetic responses to works of art [8, pp. 10-11].

It is true that this study was not a success, but what is more important is the fact that for the first time psychological data were gathered with the aim of finding out the value of paintings, in this case Holbein's Madonnas (there were some controversies about the authenticity of these paintings (cf. [9]): This is the new approach Fechner introduced to the field of aesthetics. Moreover, he pointed out three methods to be used in experimental aesthetics:

1. method of choice: subjects are asked to select the most pleasing from a number of objects

2. method of production: subjects are asked to produce some objects which they believe to be aesthetically pleasing
3. method of use: objects of interest are analyzed with respect to some characteristics which could be the reason for the appreciation of these objects—e.g., the proportion of windows, doors, books, etc.

All of these methods are still in use today and although methodology has been refined since Fechner's days [10], experimental aesthetics has kept these basic methods alive. Molnar [11, p. 6] believes that "the current themes of the psychology of art are still taste, choice, and judgement" and they are still under investigation with the help of Fechner's methods.

Of course, the Fechnerian approach has been criticized from the beginning. However, what most critics do not mention is that Fechner himself did not want to replace aesthetics "from above" by aesthetics "from below"; rather, he tried to find a basement of data that would give aesthetic theories a stock of facts. Croce [12], nevertheless, accused him of merely playing games of no value because he believed that it was not possible to make aesthetic inquiries without an analysis of concepts (= "from above") and that Fechner's experiments are nothing but a pseudoscientific waste of time.

A more justified argument against this experimental approach was given by Brentano [13]. He thought that measurement could do nothing but compare mental *products* and would never be able to say something about the mental *activity*, the psychic act itself. This is not to say that Fechner's experiments are useless, but they are the midpoint between aesthetics "from above" and "from below," because the connecting link between both is the psychic act. Brentano did not condemn Fechner's work but pointed to the active mental component in aesthetic apprehension which is represented in Fechner's thinking as "associative factor" only. This "principle of association," however, is somewhat mechanistic and has nothing of the elegant, but often misunderstood expression Brentano used in his conception of 'psychic act' [cf. 14, pp. 195-196; 15, pp. 96-97].

Fechner was completely aware of the fact that only minor problems of aesthetics could be solved by this approach and, therefore, he cautiously spoke of "Vorschule der Ästhetik" (Preparatory School of Aesthetics) and used the term "principle" instead of "law." Fechner was not cautious because he put together his more or less unsystematic thoughts and experiences with aesthetics; rather, he felt that the hypothesis of psychophysical correspondence will reach its limits with aesthetic phenomena. In other words: aesthetics may be the crucial test for psychophysics.

But, nevertheless, his method was—and still is—fascinating. Berlyne reverted to the principles of the Fechnerian approach hoping to find new results in his experimental aesthetics by using more sophisticated methods, describing stimulus properties based on information theory, using collative variables, and so on [8, 10]. Berlyne's theory, however, was criticized by showing that there are severe

problems with respect to the principle of falsifiability of the inverted u-shaped curve which is central to his thinking [16]. In short: An experimental result showing an ascending curve gives reason to conclude that one is on the left side of the inverted U. If the result is a declining curve—one is on the right, and if the result shows a straight line, one is in the plateau range of the inverted U. All these outcomes would be consistent with the theory, but the explanation will be a post-hoc one and this is not what a good theory should allow (for further criticism of Berlyne's theory, see [17, pp. 74-97]). One of the consequences of Bortz' [16] arguments is that the psychophysical approach may work only if very simple material is used as the aesthetic stimulus, and this type of material is exactly the beginning of experimental aesthetics.

THE GOLDEN SECTION HYPOTHESIS

In 1855 Zeising published his "Aesthetische Forschungen" (Aesthetic Investigations) and made a great plea for the aesthetic importance of the golden section ratio, also known as the divine proportion [18]. This mathematical relation is expressed as that specific ratio of the shorter (a) to the longer (b) of two straight lines that equals the ratio of the longer line (b) to the sum of both lines (a + b): i.e., $a : b = b : (a + b)$. The solution of this quadratic equation results numerically in .618. Thus, the proportion of line (a) is .618 times line (b) [18, p. 222].

Based on the findings that the golden ratio could be found in nature (e.g., the number of leaves in relation to the number of turns [18, p. 192], or the proportions of the human figure [18, pp. 194-195], Zeising believed that this "law of proportion" is as important as the "logical laws of truth" and that it is the "yardstick of aesthetic judgements" [18, p. 178]. As he hoped to find the "law of proportion" everywhere—in physiology, physics, chemistry, geology, and even astronomy—Fechner had doubts whether this could be true and tested Zeising's hypothesis experimentally [6, 19].

His results, however, showed that the golden section was indeed the preferred ratio. Because of this unexpected result Fechner tried to repeat this result using a different method—the method of use. He analyzed the proportions of works of art but failed to find a preference for the golden section. At this point Fechner stated that Zeising placed too much emphasis to the golden section as the "law of proportion" ("Hiernach kann ich nicht umhin, den ästhetischen Werth des goldenen Schnittes von Zeising überschätzt zu finden . . . " [6, p. 192]. Thus, the controversy had begun and even today the golden section hypothesis is "as controversial as ever" [20, p. 467].

Fechner's experiment received much attention from other researchers (see [21] for an overview) and the problem was and still is that the golden section is confirmed in some studies while in others it is not.

Lalo [22, cit after 23, p. 391] tried to repeat Fechner's experiment and found similar results (cf. Figure 1). By use of the data reported in Woodworth [23,

p. 385] we computed the rank correlation coefficient for Fechner's and Lalo's data and found $r = .39$ which does not reach the critical value of $r = .56$ ($N = 10$; $p = .05$). However, when we computed the same correlation for the data Fechner and Lalo obtained when they asked their subjects to select the *least pleasing* from a set of ten rectangles we obtained $r = .97$ ($p < .01$). Most surprisingly, this points out that the subjects' judgements of *not pleasing* rectangles are much more reliable and consistent than their judgements of pleasantness (Wohlgefälligkeit). Of course, we should be cautious to generalize these results, because there may be lots of circumstances that were not kept constant in both experiments and, therefore, our conclusion might be misleading. However, one should note that Fechner as well as Lalo included the *negative* pole of aesthetic judgements in their work—an aspect which has nearly disappeared in later studies on proportions.

Haines and Davies worked with a specially designed apparatus that their subjects had to use when producing quadrangles they believed to be pleasant [24]. The authors found a wide range of most pleasant proportions and refused to compute a mean proportion because "to use the method of averages . . . is indeed to kill the goose that lays the golden egg" [24, p. 259]. This is correct if one wants to make a statement with regard to one (and only one) proportion, namely the golden section. Yet, if one computes the mean value of their data [24, p. 256, Table 1] the golden section appears again: the mean proportion is .628 and this is close to .618 (which is the golden section).

To compare the results of Fechner's [6], Lalo's [22], and Haines and Davies' [24] studies we computed the relative frequency of choices for fifteen classes of proportion (cf. Figure 1).

As can be seen, there is some agreement between the data of [6] and [22] because both have a peak at the golden section. The results of [24], however, do not fit in this picture. And, of course, we have to state that there is a special problem with the square (or the square-like proportions Nr. 14 and 15) which "was chosen by a large enough number to suggest a second mode" [23, p. 386]. It should be noted that there are no preferences in classes one through four indicating that the range of proportions was limited by the experimental procedures and, therefore, the proportions in these classes had no chance of being the preferred ones.

Thus, in the beginning of experimental aesthetics we do not find clearcut results and even after more than a century of research the golden section hypothesis still remains unclear. This is even more so since there is a wide variety of methods in this branch (cf. [21]). To name just a few: one study [25] tried to find the most pleasing partition of a horizontal line, but the subjects were instructed to make the most pleasing *unequal* partition of the given line. Under this condition, indeed, it could be shown that people mostly used the golden section, however, this instruction prohibited that the ratio 1:1 could emerge at all.

Davis wanted to know whether there was stability over time when producing rectangles but the correlation he found between first and second drawings was .67

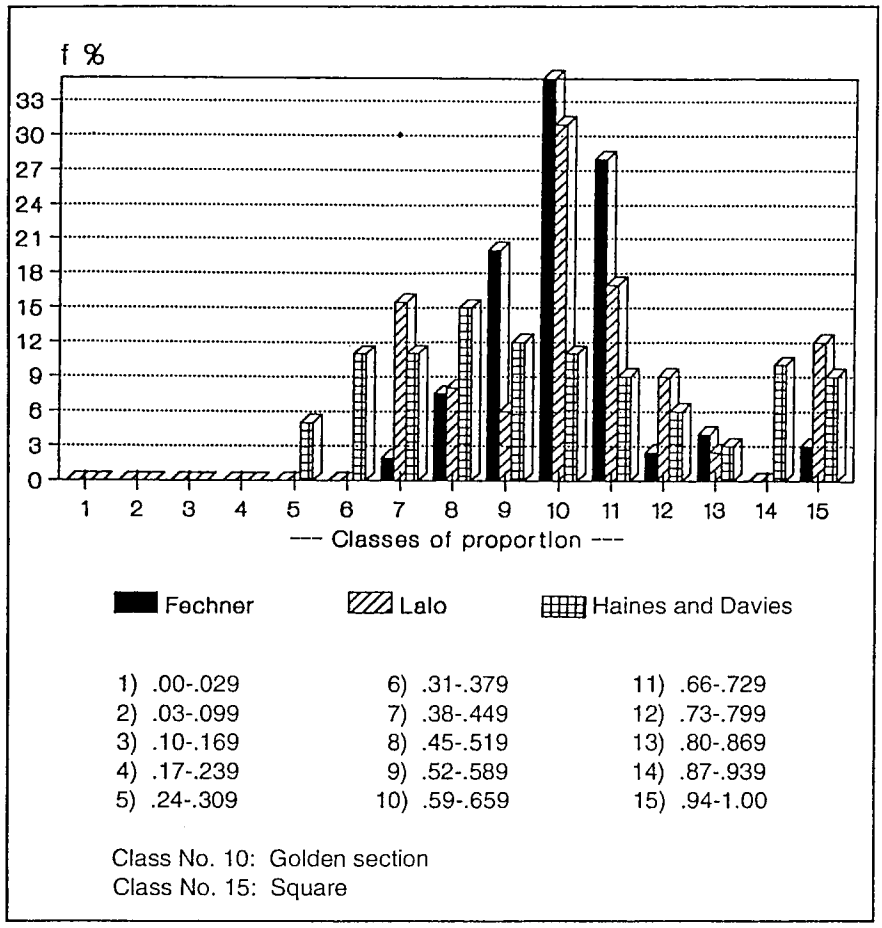

Figure 1. Distributions of relative frequencies (f%) of preferences for different proportions. In order to make the results of Fechner [6], Lalo [22], and Haines and Davies [24] comparable, percentages were computed using the original data. Thereafter, data were classified according to the fifteen classes of proportions shown above.

which—according to his opinion—was too weak to speak of a reliable aesthetic judgement [26]. Thompson investigated the developmental aspect of the aesthetic judgement and found that with increasing age the preference for the golden section increases [27]. Unfortunately, the square was not among the stimuli and thus the study cannot give any information concerning Woodworth's suggestion that the square is a serious rival to the golden section.

The criticism against the range of stimuli was the basic reason for two studies [28, 29]. According to the authors, preference for certain ratios is mainly an effect of demand characteristics of stimulus presentation. Indeed, they found that preference was influenced by the spatial position at which the rectangles were presented (at the beginning or at the end) and—most important—the golden section was *not* preferred. These studies were criticized by Benjafield [30] who argued that 1) subjects had not enough time to give a real aesthetic judgement and 2) rectangles were not controlled in size. In his own experiment he tried to overcome these disadvantages using KELLY's repertory grid technique which should enhance a more thoroughful aesthetic judgement. Under these conditions he found that the rectangle with the golden section proportion was the preferred one.

However, none of these studies could explain *why* the golden section might be of specific aesthetic value. Of course, there are some remarks about the reason for the aesthetic pleasure the golden section elicits but they are difficult to test (e.g., in [31] it is suggested that a kind of musical harmony and its being associated with visual forms were the reason why people preferred the golden section; for overview of different explanations see [8, pp. 222-232; 20, pp. 468-478; 21]).

An exception is the hypothesis of Stone and Collins [32], known as the "perimetric" hypothesis, which can be traced back to [33]. This hypothesis postulates that the rectangularly surrounded binocular visual field is about the same proportion (.665) as the golden section (.618) and, therefore, these rectangles are preferred.

This hypothesis was tested empirically [34], but the author failed to get the golden section in the rectangles his subjects drew. The second test of the perimetric hypothesis was done by Hintz and Nelson [35]. They believed that if there is a correspondence between the ratio of the perimetric rectangle of the visual field (computed according to the rules of [32]) and the ratio of the preferered rectangle, then one should find individual differences with regard to the visual field and the preferred rectangle. However, when they measured the visual field of each of their subjects and correlated the ratio of the perimetric rectangle and the preferred rectangle of that person they found $r = .279$, and this is surely not sufficient to verify the perimetric hypothesis.

Berlyne [1, p. 300] believed that there will be no simple solution with respect to preferred proportions and thought that even minute variations in instruction or experimental arrangements caused differences in results. As far as we can see, most of experimental scrutiny has been given to the stimulus conditions. The cognitive basis of the subject however, has been treated with minor attention.

In many investigations of the golden section hypothesis we find different verbal criteria given to the subjects when they were asked to make a judgement of preference. These criteria were not systematically varied and mostly used without special consideration, i.e., it was held that the criteria were more or less interchangeable (pleasing, harmonious, most preferred etc.). However, according to

Hörmann [14, p. 500] verbal communication can be used to direct the awareness of a listener and therewith giving a basis for his or her actions. If this is correct, variation of the instructions would be a method to alter the listener's mental model of the task required (cf. [36]). This aspect of communication was the reason to carry out the present experiments. As far as we know the influence of verbal criteria given in studies on proportions has never been investigated before.

Moreover, in most cases studies on the aesthetic value of the golden section were done using the method of choice. In these instances subjects are required to *select* from a given set of figures (say, rectangles). Thus, the range of proportions presented is restricted and some ratios simply cannot be chosen (cf. Figure 1). To overcome these difficulties a change in procedure seemed necessary: when subjects are asked to *produce* figures, conditions are provided allowing for every proportion to be drawn and, therefore, having a chance of being preferred over another one (cf. [20, pp. 473-474]).

Thus, it is predicted that different verbal criteria cause people to produce or select different proportions. Moreover, if the golden section hypothesis is correct, this proportion should be found more frequently when people have to work under the criterion of aesthetic attractivity (beauty).

EXPERIMENT 1

Method

Sixty-two students (at Ruhr-Universität Bochum, Germany) of both sexes and of different fields participated in the study, randomly assigned to two experimental conditions. A small booklet of eight pages each was given to each subject; every second page (80g, white, DIN-A4 format) was printed with a black straight line at the same central position on the lower third of the paper. The *g*iven *l*engths of *l*ines (GLL) were 3.1, 5.1, 7.2, and 9.2 cm. GLL was randomly assigned to control for effects of position. Moreover, every lined page was followed by a blank sheet of paper to avoid impressions of the preceding drawing that might influence the next one. Every subject received the same type of felt-tipped pen (Schwan-Stabilo-point 88) which produced thin black lines requiring low pressure of drawing. Each subject was seated at the same desk, illumination was accomplished by white neon lamps mounted at the ceiling of a windowless room. The booklets were on the desk with the shorter side being parallel to the chair of subjects. Thus, all GLLs were presented horizontally.

One half of subjects were instructed 'to make a rectangular quadrangle' (= simple instruction) by using GLL as basis (i.e., each subject produced 4 quadrangles). Remaining subjects were asked 'to make a *beautiful* rectangular quadrangle' (= beauty-instruction). Thus, the important manipulation was merely to add the word 'beautiful' to the instruction.

It should be stressed that the words of the instruction were chosen carefully to avoid confounding effects of unclear terms. Firstly, there was no opposite criterion like "ugly" in the instruction to minimize the potential effect of the beauty-criterion, i.e., to provide conditions which lower the probability of a significant difference, hence, working *against* the hypothesis. Second, the term "quadrangle" was used because "rectangle" is defined geometrically as having a pair of two parallel lines which have to be *unequal* in length and under these circumstances our subjects would have avoided to make squares (cf. [25]). Thirdly, in order to get figures which are comparable with respect to their proportions, the term 'rectangular' was added, because otherwise the computation of proportions would be meaningless. Thus, subjects had every freedom to draw the quadrangle they wanted, every proportion had the same probability of being produced.

The dependent variable was the length of the right vertical line (in cm) as drawn by the subject. If the length of this line was different from that of the left vertical line, both lengths were averaged.

Two subjects were excluded from the analysis because they drew figures which were not required; the number of subjects given above (62) is already reduced correctly.

Results

Proportions were calculated for every drawn rectangle by dividing the shorter line by the longer one (i.e., basis : vertical line or vertical line : basis). This method provides proportion values which are between zero and one, thus, making presentations more convenient.

A comparison of mean proportions by three-factorial ANOVA revealed that the type of instruction (factor 1) significantly influenced the proportions of the quadrangles our subjects drew, $F(1,247) = 7.71$, $p < .007$; means: simple instruction .64, beauty criterion .74). This is evidence that proportions which are closer to the square are more beautiful than others; this fact will be analyzed in closer detail below.

To prove the influence of GLL we included "given line length" as a second factor, but it did not reach significance ($p > .70$). This is important because it shows that regardless of GLL subjects concentrate on the *proportion* of the quadrangle dependent on the criterion given in the instruction.

As there are some studies controlling for sex (e.g., [28]) this variable constituted the third factor. As it did not reach significance either ($p > .73$), we may conclude that both sexes had the same concept of beauty in mind when drawing their quadrangles. Also, none of the interactions reached significance.

As Haines and Davies argued that "to use the method of averages . . . is indeed to kill the goose that lays the golden egg" [24, p. 259] the frequency distribution of the proportions was analyzed. We divided the range of proportions into fifteen

classes and compared the frequencies of those quadrangles drawn under the "beauty"-instruction and those drawn under the simple instruction. The result is significant ($\chi^2 = 25.55$, $df = 14$, $p < .03$, cf. Figure 2): people produce proportions differently dependent on the criterion given. Obviously, proportions lower than class 8 are much more unlikely under the beauty instruction.

Moreover, Figure 2 shows that more quadrangles belong to the square-like proportions than to the golden section. A separate comparison of frequencies in

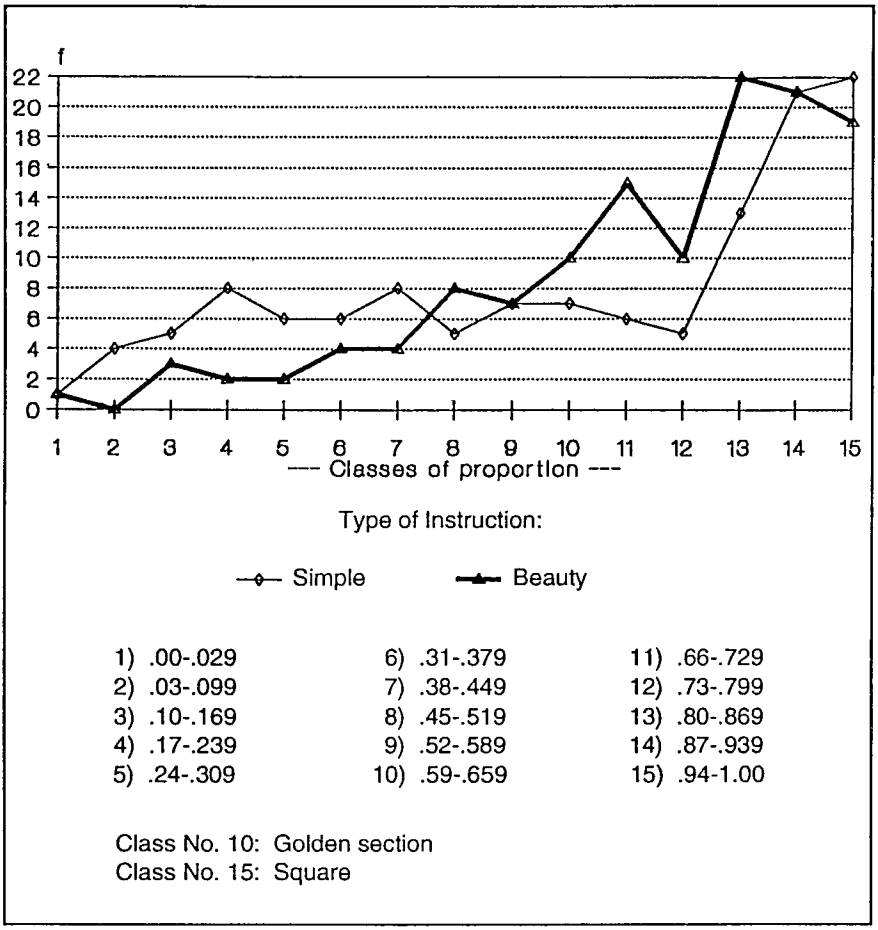

Figure 2. Frequency distributions (f) of proportions of quadrangles drawn under two experimental conditions: "simple instruction" and "beauty-instruction." According to their proportions quadrangles are classified into fifteen classes.

Class ten (golden section) and Class fifteen (square) does not approach significance (χ^2 corr = .33, $df = 1, p > .50$), thus, clearly indicating that the golden section and the square *generally* are preferred ratios, but do not differ from each other in this respect. This is good evidence that there are no special proportions which are aesthetically pleasing, but the *range* of proportions is different when the instruction is to draw beautiful quadrangles.

The next step was to try to repeat these outcomes by using a different method—against a Fechner method—i.e., the method of choice. Additionally, as we were successful to direct the awareness of our subjects by the instruction we are now interested in the question whether the drawings of the subjects in Experiment 1 fulfill a communicative function: Can different subjects decide under which instruction the quadrangles had been performed? If so, we predict that the results should show a significant correspondence between the proportions drawn under the beauty criterion and those figures selected under the same criterion.

EXPERIMENT 2

Method

Twenty subjects from different fields (at Ruhr-Universität Bochum, Germany) were asked to sort a random sample of 176 quadrangles (out of 248 drawn by the subjects of the first experiment) into two piles: one was to contain all quadrangles which they believed were produced under the instruction of beauty of the first experiment (beauty pile) and the other pile those that were done under the simple instruction (simple pile). When a subject had completed the task, data were recorded, both piles but into one, shuffled to create a new random order of quadrangles, and then presented to the next subject. Conditions of experimental setting were kept constant: i.e., subjects working in the sorting task used the same room and desk as the subjects in Experiment 1.

Results

For both piles ("beauty" and "simple") the number of proportions in each of the fifteen classes was counted (the definition of classes is given in Figure 1 and 2). By use of Kolmogorov-Smirnov statistic we tested whether the cumulated relative frequencies of proportions in each pile differed or not. Again we found a significant difference between both piles ($D_{max} = .11$; $D_{obs} = .094$; $p < .001$) indicating that even under the condition of sorting people discriminate between proportions according to the criteria given.

To check the agreement of subjects in both experiments working under the beauty instruction, we finally tested whether the frequency distribution of sorted and drawn rectangles differ or not. If they do *not* differ from each other, then (and only then) we may conclude that different subjects working with different

methods actually use the same *range* of proportions—i.e., they have the same concept of beauty and the same type of correspondence between criterion and proportion. Using the Kolmogorov-Smirnov test we found that the critical value of $D = .094$ was not reached ($D_{max} = .090$); i.e., there is no significant difference ($p > .30$). In addition, rank correlation of both frequency distributions is $r = .86$ ($p < .01$).

These results—no significant difference between sorting and drawing task with respect to the frequency distribution of proportions *and* high correlation—clearly show that the cognitive basis of subjects was the same in both experiments, which demonstrates that the instructional condition is effective in providing discriminative functions and it also points out that there is some reason to believe in the efficiency of visual communication.

DISCUSSION

First, we have to state that in both experiments results do not give any support for the golden section hypothesis. As this hypothesis predicts at least a peak of preference at .618 the facts clearly contradict this prediction. Even the method of giving some room for the deviation usually to be found in empirical research—i.e., allowing for some broader conception of the golden section by putting all ratios between .59 and .659 into one class—does not enhance support for this hypothesis. Thus, supporting the early doubts of Ogden: "The results do not show that one figure is necessarily more favored than another; and some psychologists have concluded that there are no 'favored forms' " [37, p. 210].

The picture of our results is in line with other data [20] collected with quite different material (ellipses) but also by use of a drawing method (freehand). With respect to the ratio of the axes of the ellipses it is clearly stated that "the most noticeable feature of the distribution is that there is no clear preference for any one ratio" [20, p. 484]. Additionally, Fechner (cf. [38, p. 506]) himself could not repeat his findings on rectangles by using ellipses, thus, these data seem to have some kind of stability over time.

However, results presented here show a significant and meaningful difference in the *range* of quadrangle-proportions which are believed to be beautiful (the criterion given in the instruction). Obviously, our data reveal a great variety of what people call "beautiful"—just as Berlyne [8, 10] pointed out with respect to the term "pleasing." On the other hand, it has also been confirmed that there is considerable agreement between the subjects of both experiments, and this agreement does not only refer to concepts but to the aesthetics of objects, too, clearly indicating that in the sorting task the criterion could be detected in the proportions of the quadrangles: a verbal criterion is an essential factor in investigations on proportions.

Consequently, we are not justified to speak of "*preference* for square-like proportions" because we asked our subjects to work on the criterion of *beauty*.

Although, in the Berlyne studies (using patterns as stimuli, (e.g., [39, p. 128]) it turned out repeatedly, that scales like pleasing-displeasing and like-dislike are highly correlated with ugly-beautiful (thus, constituting the 'hedonic-tone' dimension) this should be treated carefully. One may predict, that the proportions drawn under a preference instruction will not differ from those reported here but this remains to be shown. What seems to be correlated conceptually is not necessarily similar in its physical realization.

One might argue against these data that they are confounded with variables like the subject's drawing ability, motor behavior or eye movements while drawing the figures. However, the problem our subjects had to solve was so easy (making a quadrangle by using a given line as basis—i.e., to draw 3 simple straight lines) that it seems impossible that these variables could be of importance.

Another objection might be that the subjects of the sorting experiment did not select according to the *proportions* of the rectangles, instead they might have concentrated on the *quality* of figures—i.e., those drawings were selected which are better with respect to the perfection of drawing. If this were true, all better drawings must have been in the range of classes ten through fifteen only, but this is rather unlikely. Moreover this would require, that the subjects of Experiment 1 had thought of "beautiful" as a criterion for perfect drawing instead of being the criterion for proportion; again this is not confirmed by the facts.

Benjafield, Pomeroy, and Saunders found that people are able to copy visually given proportions of straight lines with considerable precision [40]. The proportions of .50 and .618 (i.e., a simple proportion and the golden section) were copied with lowest error rate and greatest number of hits, whereas other proportions (.667 and .75) had lower levels of accuracy. Unfortunately, the authors do not report data of pleasingness or the like—i.e., we do not know whether the most accurately copied ratios were evaluated positively.

Nevertheless, if the golden section is one of the most accurately delineated proportions it seems extremely unlikely that subjects of Experiment 1 could have avoided the golden section because it is more difficult to produce. Thus, if high accuracy is to be found with the golden section and a simple proportion (square and 0.50) we should have had higher frequencies in these three cases—i.e., three peaks of frequency distribution—but the results do not support such a view.

However, there are two major differences between Benjafield et al.'s [40] study and the procedure employed here: 1) subjects had to divide horizontal and vertical *lines* and 2) subjects had to *copy* the proportions presented as a standard (direct visual criterion without memory effects). In our experiment subjects had to *produce* a "beautiful" figure, i.e., they had nothing but to rely on (a) their subjective feeling of beauty, (b) probably on planning processes on how to do the quadrangle and finally (c) on the evaluation of the *quadrangle* while being created. Put differently: they had no direct visual criterion but memory was involved in the task—i.e., they had to remember the criterion.

All these arguments support Berlyne's suggestion that even minor changes of experimental conditions result in differences of the outcome. Hence, it becomes more and more difficult to decide pro or contra the aesthetic attractivity of this or that proportion.

However, Svensson reported that there is definitely a preference for the divine proportion [41]. But this contradictory result, too, can be explained by differences in the methods applied: 1) different stimulus material was used (horizontally *and* vertically presented lines of different lengths vs. always horizontally presented GLL in our experiment); 2) the task his subjects had to deal with was "to partition a line at the point where the resulting segments formed the most pleasing ratio" [41, p. 79], i.e., the subjects in his experiment were clearly informed that a *ratio* is the dependent variable. Thus, the influence of prior knowledge must be discussed.

Although Svensson reported that "none of the subjects were familiar with the meaning of the golden section" [41, p. 79] this point remains unclear as in most investigations because there are many terms for the golden section (e.g., divine proportion, harmonic partition, etc. (cf. [38]) and subjects possibly knew this proportion under another term or simply remembered that there is "something special" with respect to proportions. As most students (in [41] psychology and art students) learned at school about certain ratios which are said to be pleasant they simply may have remembered the problem. Thus, the outcome of such experiments merely may reflect that the preferred ratio is learned as a tradition of western culture (cf. Berlyne and Hatano; cit. after Berlyne, [8], pp. 229/230) and not established according to the actual impression. This again points to the efficiency of criteria given in the instructions. Therefore, the instructional texts applied in our experiments deliberately did not mention the word "ratio" or "proportion" to avoid effects of prior knowledge.

Moreover, the legends of Svensson's figures say that the subjects divided the lines "so that the partition appears harmonious and pleasant" [41, pp. 79-80]. Apart from the discrepancy of Svensson's description in the method paragraph and the legends of his figures, the terms "most pleasing" and "harmonious and pleasant" are obviously combined; consequently, subjects may have worked on their task by using a *compound criterion.* Similarly, Boselie's subjects were instructed to give "judgements of preference" [42, p. 371] by telling the experimenter "which of the two patterns he/she preferred, or thought looked best" ([42, p. 372]; similar instructions are given in many other investigations including Fechner's original study).

Additional evidence for the combined effects of verbal phrases in instructions can be found in psycholinguistics. Sichelschmidt investigated the influence of adjectives (criteria) on drawings of simple geometric forms (triangles) using descriptive adjectives like big, small, rectangular, acute-angled, obtuse-angled etc. (groβ, klein, rechtwinklig, spitzwinklig, stumpfwinklig; [36, p. 83]). Most interestingly, he found that order and type of relation of the adjectives given (e.g., subjects were instructed to draw a "great obtuse-angled triangle" or an

"obtuse-angled great triangle"; [36, pp. 84, 100]), entailed significant differences of the physical features of the figures. As these criteria were chosen such that they may serve as descriptions of physical features one can imagine that things may become much more complicated if one is interested in beauty, preference or pleasantness because all these criteria are surely not unidimensional.

Hence, it is to be expected that a "harmonious and pleasant" figure is different from a "pleasant and harmonious" one. Furthermore, the term "harmonious" most likely may have made Svensson's [41] subjects remember the harmonious division as a mathematically defined partition. Even if they did not remember the ratio exactly it seems sufficient that they remembered a proportion which is *different from equal partition* and, *therefore*, ratios close to the golden section were given preference over square-like proportions.

As we had only one criterion in our experiments and if Sichelschmidt is right, then it follows that we have no right to generalize our result to, say, preference or pleasantness etc. as already mentioned. On the other hand, this would point out that aesthetic judgements are highly sensitive to any kind of even minor variation and are, thus, a strong test for psychophysics. In order to find a solution of the problems of aesthetics it seems insufficient to describe stimulus properties physically and to rely on the verbal response of the subject; rather, it is necessary to include the cognitive and emotional disposition of the beholder as well (cf. [17, 43]). Up to now, regrettably, most researchers mainly attended to the object itself.

On this background it is not surprising that things become even more complicated if we look at other figures (not only quadrangles and lines) where the golden section should be equally effective but did not yield the expected influence. After years of careful research, Boselie [42, 44] argues that it is not the golden section that is relevant to the aesthetic pleasure people have when they prefer some rectangles. Instead, it is the detection of simple ratios (say 1:1 or 1:2) that enhances preference. However, this theory would move the problem to the question "Why is the detection of simple ratios pleasing?" The important fact is, that it depends on the subject whether he or she detects the simple ratio. This is a creative achievement of the beholder, not an effect of the stimulus per se.

Prinz pointed out that it is not the stimulus in itself that is relevant for behavior; rather its being related to stored information, to human—and probably not only human—systems of reference that constitute meaning and this meaning is relevant for behavior [15, p. 33]. Perhaps, Boselie and Leeuwenberg's line of thinking is able to cast some light on the factors of this process by their "means-effect" hypothesis referring basically to the unity in variety conviction but going a step further [45]. As the unity in variety hypothesis is not able to explain the beauty of the square (because there is no variety) the authors claim for the importance of the success of the beholder in detecting hidden regularities in patterns. If the means employed have maximum effect then aesthetic pleasure will arise. In detail: aesthetic feelings arise from the " 'formal' quality of contrast between means and

effect" [45, p. 8]. Thus, the square may be pleasing because its regularity (sides of equal length) is easy to detect.

After 130 years of empirical research on the golden section hypothesis (from 1865 to 1995), the only conclusion to be made is that the results are still contradictory and we should perhaps give up the psychophysical hope that in aesthetics, too, relations can be found between stimulus properties and aesthetic judgements which could be termed "law" or "principle."

Currently, a tendency is to be seen to extend the validity of the golden section as a proportion with special relevance into psychological areas like personality studies etc. Even if significant results are obtained (e.g., the evaluation of personal acquaintances show a 62% preference for positive dimensions of evaluation; see Rigdon and Epting [46]) this line should be treated with care because it is at least too early to base the assumptions of such investigations on the aesthetic importance of visual material made according to the divine proportion—the hypothesis is still nothing but a hypothesis and not a safe ground to build systems of lifestyle-event partitioning.

Finally, emphasis should be given to the possible application of research on the golden section. If one thinks of the golden section as a basis of design—as Corbusier did in the conception of the Modulor or Hambidge (cf. [21])—the problem is even more serious because designers relying on the beauty of the divine proportion may be completely wrong, say, in the built environment. On the other hand it might be too early to conclude "that the golden section hypothesis should die a natural death" as Plug proposes [20, p. 486]. The possibility still remains that—under certain circumstances (some of which have been worked out by Boselie [42, 44]) or for special purposes—the golden section indeed might be of highest aesthetic value. However, it should be found out what kind of circumstances and what kind of purpose call for the divine proportions.

REFERENCES

1. D. E. Berlyne, *Konflikt, Erregung, Neugier. Zur Psychologie der kognitiven Motivation*, Klett, Stuttgart, 1974.
2. W. Platz, Fechner's Leben und Wirken, in *Das Büchlein vom Leben nach dem Tode*, G. T. Fechner (ed.), Strecker und Schröder, Stuttgart, 1922.
3. L. Oken, Über das Universum, in *Gesammelte Schriften*, J. Schuster (ed.), Keiper, Berlin, 1939; originally published 1808.
4. C. G. Allesch, *Geschichte der psychologischen Ästhetik*, Hogrefe, Göttingen, 1987.
5. N. Bischof, Erkenntnistheoretische Grundlagenprobleme der Wahrnehmungspsychologie, in *Allgemeine Psychologie. Handbuch der Psychologie I, 1*, W. Metzger (ed.), Hogrefe, Göttingen, 1974.
6. G. T. Fechner (1876/1925/1978), *Vorschule der Ästhetik*, Olms, Hildesheim, Olms, 1978, originally published 1876.
7. F. Molnar, Experimental Aesthetics or the Science of Art, *Leonardo, 7*, pp. 23-26, 1974.

8. D. E. Berlyne, *Aesthetics and Psychobiology*, Appleton-Century-Crofts, New York, 1971.
9. P. H. Marshal, J. Worthen, L. Brant, J. Shrader, D. Kahlstorf, and C. Pickeral, Fechner Redux: *A Comparison of the Holbein Madonnas*, unpublished manuscript, Texas Tech University, 1994.
10. D. E. Berlyne, The New Experimental Aesthetics, in *Studies in the New Experimental Aesthetics*, D. E. Berlyne (ed.), Wiley, New York, pp. 1-25, 1974.
11. F. Molnar, *Fechner's Project a Hundred Years Later*, paper presented at the International Conference on Psychology and the Arts, Cardiff, Wales, Great Britain, 1983.
12. B. Croce, *Ästhetik als Wissenschaft des Ausdrucks und allgemeine Linguistik*, Seemann, Leipzig, 1905.
13. F. Brentano, *Psychologie vom empirischen Standpunkt* (3 Vol.), Meiner, Hamburg, 1955-1968, originally published 1874.
14. H. Hörmann, *Meinen und Verstehen*, Suhrkamp, Frankfurt, 1976.
15. W. Prinz, *Wahrnehmungs-und Tätigkeitssteuerung*, Springer, Berlin, 1983.
16. J. Bortz, Psychologische Ästhetikforschung—Bestandsaufnahme und Kritik, *Psychologische Beiträge, 20*, pp. 481-508, 1978.
17. H. Höge, *Emotionale Grundlagen ästhetischen Urteilens*, Lang, Frankfurt, 1984.
18. A. Zeising, *Aesthetische Forschungen*, Medinger, Frankfurt, 1855.
19. G. T. Fechner, Über die Frage des goldenen Schnittes, *Archiv für die zeichnenden Künste, 11*, pp. 100-112, 1865.
20. C. Plug, The Golden Section Hypothesis, *American Journal of Psychology, 93*, pp. 467-487, 1980.
21. H. J. McWhinnie, *A Review of the Research on the Golden Section Hypothesis in Art and Design*, paper presented at the Honas Seminar on the Golden Section, 1990.
22. C. Lalo, *L'esthetique expèrimentale contemporaine*, Alcan, Paris, 1908.
23. R. S. Woodworth, *Experimental Psychology*, Methuen, London, 1938.
24. T. H. Haines and S. E. Davies, Psychology of Aesthetic Reaction to Rectangular Forms, *Psychological Review, 11*, pp. 249-281, 1904.
25. R. P. Angier, The Aesthetics of Unequal Division, *Psychological Review Monograph Supplement, 4*, pp. 541-561, 1903.
26. F. C. Davis, Aesthetic Proportion, *American Journal of Psychology, 45*, pp. 298-302, 1933.
27. G. C. Thompson, The Effect of Chronological Age on Aesthetic Preference for Rectangles of Different Proportions, *Journal of Psychology, 36*, pp. 50-58, 1946.
28. M. Godkewitsch, The "Golden Section": An Artifact of Stimulus Range and Measure of Preference, *American Journal of Psychology, 87*, pp. 269-277, 1974.
29. J. Piehl, The "Golden Section": An Artifact of Stimulus Range and Demand Characteristics, *Perceptual & Motor Skills, 43*, pp. 47-50, 1976.
30. J. Benjafield, The "Golden" Rectangle: Some New Data, *American Journal of Psychology, 89*, pp. 737-743, 1976.
31. E. H. Huntley, *The Divine Proportion*, Dover, New York, 1970.
32. L. A. Stone and L. G. Collins, The Golden Section Revisited: A Perimetric Explanation, *American Journal of Psychology, 78*, pp. 503-506, 1965.
33. M. Borissavlievitch, *The Golden Number*, Tiranti, London, 1970.

34. H. R. Schiffman, Golden Section: Preferred Figural Orientation, *Perception & Psychophysics, 1*, pp. 193-194, 1966.
35. J. M. Hintz and T. M. Nelson, Golden Section: Reassessment of the Perimetric Hypothesis, *American Journal of Psychology, 83*, pp. 126-129, 1970.
36. L. Sichelschmidt, *Adjektivfolgen. Eine Untersuchung zum Verstehen komplexer Nominalphrasen*, Westdeutscher Verlag, Opladen, 1989.
37. R. M. Ogden, Naive Geometry in the Psychology of Art, *American Journal of Psychology, 49*, pp. 198-216, 1937.
38. I. C. McManus, The Aesthetics of Simple Figures, *British Journal of Psychology, 71*, pp. 505-524, 1980.
39. D. E. Berlyne, Verbal and Exploratory Responses to Visual Patterns Varying in Uncertainty and in Redundancy, in *Studies in the New Experimental Aesthetics*, D. E. Berlyne (ed.), Wiley, New York, 1974.
40. J. Benjafield, E. Pomeroy, and M. A. Saunders, The Golden Section and the Accuracy with which Proportions are Drawn, *Canadian Journal of Psychology, 34*, pp. 253-256, 1980.
41. L. T. Svensson, Note on the Golden Section, *Scandinavian Journal of Psychology, 18*, pp. 79-80, 1977.
42. F. Boselie, The Aesthetic Attractivity of the Golden Section, *Psychological Research, 45*, pp. 367-375, 1984.
43. H. Höge, The Emotional Impact on Aesthetic Judgements: An Experimental Investigation of a Time-honored Hypothesis, *Visual Arts Research, 10*, pp. 37-48, 1984.
44. F. Boselie, Complex and Simple Proportions and the Aesthetic Attractivity of Visual Patterns, *Perception, 13*, pp. 91-96, 1984.
45. F. Boselie and E. Leeuwenberg, Birkhoff Revisited: Beauty as a Function of Effect and Means, *American Journal of Psychology, 98*, pp. 1-39, 1985.
46. M. A. Rigdon and F. R. Epting, A Test of the Golden Section Hypothesis with Elicited Constructs, *Journal of Personality and Social Psychology, 43*, pp. 1080-1087, 1982.

Direct reprint requests to:

Dr. Holger Höge
University of Oldenburg
Department of Psychology
Institute for Research into
Person-Environment-Relations
Postbox 2503,
D-26111 Oldenburg
Federal Republic of Germany