

TRIADIC COMPARISONS OF MUSICAL INTERVALS¹

By W. J. M. LEVELT, J. P. VAN DE GEER and R. PLOMP

Institute for Perception RVO-TNO, Soesterberg, The Netherlands

An analysis is made of the perception of musical intervals. Two kinds of stimuli were used; intervals consisting of two simultaneous simple tones (sinusoids) and intervals consisting of two simultaneous complex tones (fundamental plus harmonics). Subjects judged the stimuli by the method of triadic comparisons in an incomplete balanced design. Multidimensional analysis was performed according to Kruskal's MDSCAL program.

The following results were obtained: (a) In both the space of simple-tone intervals (*S*) and the space of complex-tone intervals (*C*) musical intervals are ordered according to their width (frequency difference between fundamental tones). This ordering appears to follow a scale which is bowed upwards in the centre; extremely narrow and extremely wide intervals are therefore more similar than would be expected on the basis of their width alone. (b) In addition, the intervals in *C* are ordered along a dimension which is related to the complexity of the frequency ratio of the fundamental tones. This is much less true for *S*. (c) In *S*, on the other hand, intervals are differentiated on a dimension which is interpreted as indicating their resemblance to certain normative reference intervals.

Attention is given to a number of methodological issues.

1. INTRODUCTION

Earlier work by the present authors (Van de Geer, Levelt and Plomp, 1962) on the perception of the musical consonance of tone intervals showed that 'consonance' for our subjects was virtually identical with evaluation. Intervals which were judged 'consonant' were also considered 'beautiful' and 'euphonious', whereas intervals called 'dissonant' were also described as 'ugly' and 'noneuphonious'. Apart from this evaluative dimension, other dimensions were also found present in what we may call the 'psychological space' of musical intervals.

The main reason for supplementing the previous work is that we wanted to obtain greater insight into the nature of the psychological space without using verbal scaling techniques; i.e., without using instructions which force the subject to judge an interval in terms of some prescribed verbal category. Such categories may well be different from those a subject would have spontaneously adopted for differentiating his auditory impressions. Therefore, we decided to

¹This study is part of a research project carried out under a grant from the Netherlands Organization for the Advancement of Pure Research (ZWO). The authors are indebted to Dr. J. B. Kruskal who sent a copy of the card deck of MDSCAL, to the Central Computer Institute of the University of Leyden for computer assistance, and to Mr. A. M. Mimpfen who carried out the experiments.

use a technique in which both the nature and the number of psychological dimensions were not determined beforehand. The most elegant method for this purpose is the method of triadic comparison. The subject is presented with three stimuli and only has to decide which two stimuli are most alike and which two are least alike, without having to verbalize in what respect the stimuli are similar or dissimilar.

The use of this method raised a number of methodological issues, both with respect to design and analysis, which are of sufficient general interest to justify their extensive treatment in this paper.

2. EXPERIMENTS

Stimuli The stimuli were tone intervals, made up of two simultaneously heard tones with a fixed ratio between their frequencies. Fifteen stimuli were used: the twelve musical intervals within the octave; and in addition two wider intervals (4 : 9 and 2 : 5) and one narrow interval between minor and major second (11 : 12). The frequency ratios for these 15 intervals are given in Table 1. This table also specifies the frequencies of the fundamental tones, and the frequency difference for each interval. The frequencies were chosen in such a way that for each interval the mean value was 500 cycles per second (cps). This was done because in the former experiment it was found that pitch determines a considerable proportion of the variance in judgements about the intervals. By eliminating pitch as a source of variation we hoped other pertinent dimensions would become more salient.

TABLE 1. SPECIFICATION OF THE 15 STIMULI. FREQUENCIES ARE GIVEN IN CPS

<i>Stimulus number</i>	<i>Frequency ratio</i>	<i>Frequencies</i>	<i>Frequency difference (d)</i>	<i>Log first common partial (a)</i>
1	1 : 2	333 : 666	333	2.82
2	2 : 3	400 : 600	200	3.08
3	3 : 4	429 : 572	143	3.23
4	2 : 5	286 : 715	429	3.16
5	3 : 5	375 : 625	250	3.27
6	4 : 5	444 : 555	111	3.35
7	5 : 6	455 : 546	91	3.44
8	4 : 7	364 : 637	263	3.41
9	5 : 7	415 : 581	166	3.46
10	5 : 8	385 : 616	231	3.49
11	4 : 9	308 : 693	385	3.44
12	8 : 9	472 : 531	59	3.63
13	11 : 12	473 : 516	43	3.75
14	8 : 15	344 : 645	301	3.71
15	15 : 16	480 : 512	32	3.89

Two different sets of 15 stimuli were used. Either the component tones were both simple tones (sinusoids), or they were both complex tones (in which the harmonics up to 4000 cps were included, all with equal loudness). For convenience these intervals will be referred to as 'simple-tone intervals' and 'complex-tone intervals', respectively.

Apparatus and presentation The stimuli were produced by sinewave oscillators with 600 Ω output impedance. In the case of the simple-tone intervals the outputs were interconnected through resistors of 12,000 Ω to avoid coupling. For the complex-tone intervals

the sinewaves were used to initiate short periodic impulses (duration 0.2 msec) and the low-impedance outputs of these pulseformers were interconnected through resistors of 120,000 Ω . A low-pass filter with a cut-off frequency of 4000 cps was used, so the complex tones consisted of all harmonics up to 4000 cps with equal amplitude. The stimuli were reproduced by a loudspeaker in front of the subject at a distance of about 1.25 m. The sound-pressure level at the position of the subject's head was 55–60 dB. The subject was seated in a soundproof room with sound-absorbing walls.

S was provided with three push-buttons located at the angles of an equilateral triangle. Each button corresponded to a particular tone interval, so six different oscillators were used. Between the presentations the experimenter, seated in another room, adjusted the frequencies of these oscillators and their ratio very accurately by means of an electronic counter and an XY-oscilloscope (making use of Lissajous' figures).

Before a test started, written instructions were presented to the subject. He was informed that, by pressing each of the three push-buttons, a different tone interval could be heard. His task was to listen to these tone intervals successively and to decide which two intervals were most similar and which two were least similar. S noted his responses on a prepared form which showed a diagram representing the positions of the push-buttons. In the test session S first made two preliminary comparisons in order to become familiarized with the task.

Experimental design It is difficult to develop a proper design for an experiment of this type. The major problem is that for a complete experiment S would have to judge 455 triads (all possible combinations of three stimuli out of the set of 15). This task is much too time-consuming to be practicable.

We therefore had to develop an incomplete balanced design which would provide for blocks of triads which were of a suitable size for one subject. The construction of these blocks was governed by the following requirements.

- (a) In order to have a block properly balanced, all 105 pairs of stimuli must occur once in it. These 105 pairs were arranged in 35 triads with each stimulus present in 7 different triads.
- (b) In order to have the appearance of the same stimulus properly spaced, each block was made up of a sequence of 7 subsets of 5 triads with each stimulus appearing once in each subset.
- (c) The relative similarity between any two stimuli within a triad will be dependent upon the third stimulus. Since in any block a pair of stimuli occurs only once, variation in the third stimulus can be only achieved by giving it different values in different blocks. In other words, blocks had to have no triads in common.

Table 2 shows the solution which was finally adopted¹. It consists of four blocks. The authors do not know whether a solution with more than four blocks is theoretically possible. However, it was considered that with four blocks the third requirement was sufficiently met, and that a search for other suitable blocks was not necessary.

¹Considerable help in finding this solution was given by Ball's treatment of a similar problem: 15 malicious boarding school children go for a walk—in rows of three—on each of the seven afternoons of the week; how can one achieve different combinations in the rows for every day? (Ball, 1939.)

TABLE 2. BALANCED INCOMPLETE SOLUTION OF THE TRIADIC DESIGN. NUMBERS REFER TO THE FIRST COLUMN OF TABLE 1. EACH ROW OF THREE NUMBERS CONSTITUTES ONE TRIAD

	1	2	3	4	5	6	7
A	9 10 13	5 8 10	6 10 11	1 2 10	10 14 15	4 10 7	3 12 10
	6 7 2	12 15 2	13 15 4	8 4 3	11 3 9	1 9 5	14 5 6
	3 1 15	9 4 14	5 7 3	6 12 9	2 13 5	15 8 6	4 11 2
	5 12 4	6 3 13	2 8 9	15 11 5	4 1 6	3 14 2	9 7 15
	8 11 14	11 7 1	1 14 12	14 7 13	7 12 8	12 13 11	13 8 1
B	13 12 1	9 14 12	4 12 8	15 11 12	12 3 6	5 12 10	7 2 12
	10 4 11	2 6 11	1 6 5	14 5 7	8 7 13	15 13 9	3 9 4
	7 15 6	13 5 3	9 10 7	4 2 13	11 1 9	6 14 4	5 8 11
	9 2 5	4 7 1	11 14 13	6 8 9	5 15 4	7 3 11	13 10 6
	14 8 3	8 10 15	15 3 2	3 10 1	10 2 14	2 1 8	1 14 15
C	2 11 14	10 12 11	15 11 13	7 4 11	11 9 8	1 11 3	6 5 11
	3 15 4	5 8 4	14 8 1	12 1 6	13 6 2	7 2 10	9 10 15
	6 7 8	2 1 9	10 3 6	15 5 2	4 14 10	8 12 15	1 13 4
	10 5 1	15 6 14	4 12 2	8 13 10	1 7 15	6 9 4	2 3 8
	12 13 9	13 3 7	7 9 5	9 3 14	3 5 12	5 14 13	14 12 7
D	15 7 8	14 10 7	1 7 9	3 6 7	7 5 11	13 2 7	12 7 4
	2 1 6	4 11 6	8 11 13	10 13 12	9 15 12	3 15 14	5 14 1
	12 3 11	15 13 5	14 2 12	1 4 15	6 8 14	11 10 1	13 9 6
	14 4 13	1 12 8	6 10 15	11 9 14	13 3 1	12 5 6	15 2 11
	10 9 5	9 2 3	3 5 4	5 2 8	2 4 10	4 8 9	8 10 3

Each of the four blocks was presented to four different Ss. They all judged one block of stimuli twice; two Ss first judged simple-tone intervals and one or two weeks later complex-tone intervals, the other two Ss listened to the intervals in the reverse order. Furthermore, for two Ss the stimuli given in Table 2 were presented in the order from left to right, and for the other two Ss in the reverse order. As four blocks were used, a total of 16 Ss took part in the experiment.

3. RESULTS

The similarity index For each of the 105 stimulus pairs a similarity index was calculated. Essentially, this index is just a count of how often a pair is judged more similar than other pairs. For instance, if for triad (i, j, k) pair (i, j) is judged most similar and pair (i, k) least similar, this would contribute 2 points to the index of pair (i, j) , 1 point to the index of pair (j, k) and zero to the index of pair (i, k) . The maximum value of the index is 32 (since we have pair (i, j) once in each of the four blocks, each block being repeated over four subjects); the minimum value is zero.

The 15×15 similarity matrix is reproduced in Table 3, both for simple-tone intervals (right upper triangle) and for complex-tone intervals (left lower triangle).

TABLE 3. SIMILARITY INDICES FOR SIMPLE-TONE INTERVALS (RIGHT UPPER TRIANGLE) AND FOR COMPLEX-TONE INTERVALS (LEFT LOWER TRIANGLE)

<i>Stimulus number</i>	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1		14	15	25	73	14	9	22	9	12	29	9	3	26	8
2	21		21	16	25	18	13	13	22	27	10	14	10	18	9
3	18	18		12	24	25	22	14	27	17	12	13	10	10	15
4	21	20	18		18	7	10	14	10	18	30	6	14	13	9
5	12	21	25	17		20	15	30	13	24	17	12	11	18	12
6	11	19	24	12	25		28	18	24	21	8	23	17	17	14
7	9	15	22	16	18	26		14	25	20	8	28	22	10	19
8	10	18	13	20	26	12	20		16	27	20	15	11	26	7
9	6	16	31	8	15	25	21	20		22	9	14	8	19	8
10	9	20	17	12	26	20	24	25	24		10	13	7	18	6
11	11	9	15	22	17	12	10	21	13	24		6	14	13	9
12	9	10	17	8	7	23	21	21	16	11	13		32	7	29
13	5	5	13	10	9	15	18	10	6	8	9	30		10	32
14	14	8	13	10	17	16	15	27	15	17	21	20	22		7
15	11	6	13	9	11	15	16	4	8	10	13	28	23	19	

Spatial configurations The most suitable technique for the further analysis of such similarity matrices is undoubtedly the Kruskal–Shepard multi-dimensional scaling program (MDSCAL). The merits of this technique are described in the original papers (Kruskal, 1964 a, b); therefore, three short remarks suffice to introduce it here.

- (a) The rationale of MDSCAL is to produce a spatial configuration for a number of stimuli in such a way that the interstimulus distances in the space have a monotonic inverse relation to the similarity indices. No further assumptions about the nature of this relation are necessary (as, for instance, the strong assumptions involving the normal probability integral which are used in classical paired-comparison analysis) because the amount of constraint in the similarity data is sufficient to produce a unique metric solution. In this sense MDSCAL gives a non-parametric solution.
- (b) MDSCAL will produce the best-fitting configuration for each number of dimensions smaller than the number of stimuli.
- (c) Which of these configurations will be adopted as the final solution is a rational decision which balances the requirement of a minimum number of dimensions, against the degree of ‘stress’ in the solution. ‘Stress’ is a measure of the extent to which the monotonic relation is violated for each degree of dimensionality. In fact, it is a residual sum of squares which, as in analysis of variance techniques, can be expressed as a percentage of the total variance.

The two similarity matrices, given in Table 3, were separately analysed by MDSCAL (on an IBM 7090, with Euclidean distance functions and primary treatment for ties; cf. Kruskal, 1964 a). The analysis started from a 10-dimensional solution and worked backwards to 1 dimension. The stress for

TABLE 4. STRESS OF MDSCAL SOLUTIONS (IN PERCENTAGES) AS DEPENDENT UPON DIMENSIONALITY (COMPUTATION STOPPED WHEN STRESS OF 5 PER CENT OR LESS WAS REACHED)

<i>Number of dimensions</i>	10-5	4	3	2	1
<i>Simple-tone intervals</i>	≤ 5.0	7.1	11.1	14.9	43.5
<i>Complex-tone intervals</i>	≤ 5.0	8.1	11.5	20.9	33.3

each dimensionality is given in Table 4. In view of the fact that an incomplete design was used, and also that different subjects judged different triads, a 'fair' stress (in the sense used by Kruskal) was considered to be acceptable. Therefore, for both matrices the 3-dimensional solution was accepted. Table 5 gives coordinates of the configurations, both for simple-tone intervals (coordinate matrix **S**) and for complex-tone intervals (matrix **C**).

TABLE 5. THREE-DIMENSIONAL MDSCAL SOLUTIONS FOR SIMPLE-TONE INTERVALS (S) AND FOR COMPLEX-TONE INTERVALS (C)

<i>Simple-tone intervals (S)</i>			<i>Stimulus number</i>	<i>Complex-tone intervals (C)</i>		
S_1	S_2	S_3		C_1	C_2	C_3
-0.734	0.096	-0.731	1	-1.261	-0.234	-0.122
0.714	0.596	0.032	2	-0.857	0.223	0.620
-0.292	0.088	0.900	3	-0.294	-0.358	0.652
0.041	0.133	-1.105	4	-0.924	0.453	-0.364
0.061	0.848	-0.045	5	0.098	0.607	0.418
-0.314	-0.389	0.630	6	0.168	-0.658	0.471
0.246	-0.486	0.701	7	0.520	-0.400	0.346
-0.365	0.692	-0.075	8	0.364	0.738	-0.092
0.118	0.266	0.923	9	0.459	-0.142	0.972
0.460	0.745	0.272	10	0.387	0.578	0.539
-0.496	0.042	-1.089	11	-0.099	0.943	-0.517
0.458	-0.839	0.223	12	0.469	-0.702	-0.404
0.382	-1.111	-0.160	13	0.435	-0.623	-0.954
-0.868	0.396	-0.328	14	0.585	0.368	-0.641
0.590	-1.077	-0.149	15	0.146	-0.793	-0.923

Further analysis of the spatial configurations The further interpretation of these two configurations will be attempted in two steps. First we shall investigate whether there are 'common dimensions' in both spaces. We mean by this that a dimension in the simple-tone interval space may be associated to a dimension in the complex-tone interval space by virtue of a high degree of correlation between the projections of corresponding interval points on these dimensions. This would mean simply that the interpretation of a common dimension can be the same for both simple-tone and complex-tone intervals.

Secondly, we may find specific dimensions; i.e., dimensions which are characteristic for only one space. The interpretation of such a dimension

should reveal something about the specific nature of simple-tone intervals as opposed to complex-tone intervals.

(a) *Matching the two configurations* The first step, of course, is to find out whether there are common dimensions at all. Geometrically this means that we want to superimpose the two spaces on one set of coordinates in such a way that by suitable rotation of the axes we achieve maximum similarity between the two configurations. Imagine the two configurations pictured in one three-dimensional space (with identical centres of gravity); we then shall have a good fit between the two configurations when points from one configuration are in the immediate neighbourhood of corresponding points from the other configuration. A criterion for fit might be that the sum of squares of distances between corresponding points is minimum.

It can be shown that this criterion is the same as that of achieving maximum covariance between projections on corresponding axes.

In terms of matrix algebra the solution proceeds as follows. Given the two matrices \mathbf{S} and \mathbf{C} of Table 5, the final solution will imply that we transform \mathbf{S} and \mathbf{C} by means of transformation matrices \mathbf{X} and \mathbf{Y} , so that $\mathbf{SX} = \mathbf{U}$ and $\mathbf{CY} = \mathbf{V}$. A condition is that $\mathbf{X}'\mathbf{X}$ and $\mathbf{Y}'\mathbf{Y}$ are both unit matrices (since \mathbf{X} and \mathbf{Y} specify orthogonal rotations). Diagonal elements of $\mathbf{U}'\mathbf{V}$ (after division by 15: the number of stimuli) represent covariances. Let us consider the first diagonal element $u_1'v_1 = x_1'\mathbf{S}'\mathbf{C}y_1$. To find a maximum for this element, we use the method of undetermined multipliers, i.e. we take partial derivatives of $\mathbf{F} = x_1'\mathbf{S}'\mathbf{C}y_1 - \frac{1}{2}\mu_1 x_1'x_1 - \frac{1}{2}\lambda_1 y_1'y_1$ and set these equal to zero.

This produces two equations:

$$\delta\mathbf{F}/\delta x_1' = \mathbf{S}'\mathbf{C}y_1 - \mu_1 x_1 \quad (1)$$

$$\delta\mathbf{F}/\delta y_1' = \mathbf{C}'\mathbf{S}x_1 - \lambda_1 y_1. \quad (2)$$

Multiplication of (1) by x_1' gives: $x_1'\mathbf{S}'\mathbf{C}y_1 = \mu_1 x_1'x_1 = \mu_1$

and of (2) by y_1' gives: $y_1'\mathbf{C}'\mathbf{S}x_1 = \lambda_1 y_1'y_1 = \lambda_1$.

Since $x_1'\mathbf{S}'\mathbf{C}y_1 = y_1'\mathbf{C}'\mathbf{S}x_1$ we have $\mu_1 = \lambda_1$.

Multiplication of eqn. (1) by $\mathbf{C}'\mathbf{S}$ gives $\mathbf{C}'\mathbf{S}\mathbf{S}'\mathbf{C}y_1 = \mu_1 \mathbf{C}'\mathbf{S}x_1$.

From (2) we know that $\mathbf{C}'\mathbf{S}x_1 = \mu_1 y_1$.

It follows that $\mathbf{C}'\mathbf{S}\mathbf{S}'\mathbf{C}y_1 = \mu_1^2 y_1$.

Similarly it can be shown that $\mathbf{S}'\mathbf{C}\mathbf{C}'\mathbf{S}x_1 = \mu_1^2 x_1$.

Therefore, x_1 and y_1 are eigenvalues of $\mathbf{S}'\mathbf{C}\mathbf{C}'\mathbf{S}$ and $\mathbf{C}'\mathbf{S}\mathbf{S}'\mathbf{C}$, respectively. The eigenvalues of these two matrices are identical (equal to squared covariances, after division by 15^2). The matrices $\mathbf{S}'\mathbf{C}\mathbf{C}'\mathbf{S}$ and $\mathbf{C}'\mathbf{S}\mathbf{S}'\mathbf{C}$, of course, in our case will have three eigenvalues with three corresponding eigenvectors. Since the matrices are symmetric the eigenvectors will be orthogonal. Also, since an eigenvector is undermined up to a constant of proportionality it can always be written so that the sum of its squared elements is unity. The matrices of eigenvectors, \mathbf{X} and \mathbf{Y} , then will be orthogonal transformation matrices since they fulfill the requirement that $\mathbf{X}'\mathbf{X}$ and $\mathbf{Y}'\mathbf{Y}$ should be unit matrices.

$\mathbf{U}'\mathbf{V}$ will be a diagonal matrix which can be arranged to have decreasing order of magnitude of its diagonal elements. The solution implies that u_i and v_i will be 'corresponding' dimensions, whereas u_i and v_j ($i \neq j$) are independent. The amount of correspondence between u_i and v_i is indicated by $u_i'v_i/15 = \mu_i/15$ (the covariance); one may also compute correlations.

The actual solution is given in Tables 6 and 7. Table 6 specifies transformation matrices \mathbf{X} and \mathbf{Y} and it also gives the corresponding eigenvalues.

TABLE 6. TRANSFORMATION MATRICES \mathbf{X} AND \mathbf{Y} TO BE APPLIED TO \mathbf{S} AND \mathbf{C} IN ORDER TO GIVE MATCHED CONFIGURATIONS $\mathbf{SC}=\mathbf{U}$ AND $\mathbf{CY}=\mathbf{V}$. EIGENVALUES OF $\mathbf{S'CC'S}$ AND $\mathbf{C'SS'C}$ CORRESPONDING TO EIGENVECTORS WHICH ARE COLUMNS OF \mathbf{X} AND \mathbf{Y}

x_1	x_2	x_3	y_1	y_2	y_3	Eigenvalues
-0.229	0.099	0.969	-0.441	0.291	-0.849	33.14
0.824	0.550	0.136	0.892	0.040	-0.450	24.74
-0.519	0.830	-0.206	0.097	0.956	0.277	0.48

Table 7 gives the final matrices \mathbf{U} (for simple-tone intervals) and \mathbf{V} (for complex-tone intervals).

In order to get a better appreciation of the degree of correspondence between the respective dimensions, correlations have been calculated. They are $r(u_1, v_1)=0.935$, $r(u_2, v_2)=0.944$, and $r(u_3, v_2)=0.064$ (see Table 10). This shows that u_1 and v_1 are practically identical; the same applies for u_2 and v_2 . However, u_3 and v_3 , are specific dimensions in the two configurations.

TABLE 7. MATCHED CONFIGURATIONS \mathbf{U} (SIMPLE-TONE INTERVALS) AND \mathbf{V} (COMPLEX-TONE INTERVALS)

$\mathbf{SX}=\mathbf{U}$			Stimulus number	$\mathbf{CY}=\mathbf{V}$		
u_1	u_2	u_3		v_1	v_2	v_3
0.624	-0.627	-0.548	1	0.336	-0.493	1.142
0.313	0.425	0.766	2	0.637	0.352	0.799
-0.329	0.766	-0.456	3	-0.126	0.523	0.591
0.674	-0.840	0.285	4	0.776	-0.599	0.480
0.708	0.435	0.185	5	0.625	0.395	-0.074
-0.557	0.278	-0.487	6	-0.615	0.473	0.284
-0.820	0.339	0.028	7	-0.553	0.466	-0.166
0.692	0.282	-0.244	8	0.489	0.047	-0.667
-0.287	0.924	-0.040	9	-0.253	1.057	-0.057
0.369	0.681	0.491	10	0.397	0.651	-0.439
0.712	-0.930	-0.251	11	0.835	-0.485	-0.484
-0.911	-0.231	0.284	12	-0.872	-0.278	-0.194
-0.924	-0.706	0.252	13	-0.840	-0.810	-0.353
0.693	-0.140	-0.720	14	0.008	-0.428	-0.840
-0.943	-0.658	0.456	15	-0.861	-0.872	-0.023

(b) *Interpretation of common dimensions.* Figure 1 shows projections of the interval points on the superimposed planes (u_1, u_2) and (v_1, v_2) . The joint coordinates have been re-baptized η and ω . The similarity between the two configurations is apparent from the relatively short distances between corresponding stimulus points.

The figure shows that the two configurations have a horseshoe-like structure in common. Although the linear correlations $r(u_1, u_2)$ and $r(v_1, v_2)$ are both low (-0.049 and 0.099 , respectively, see Table 10) the presence of the horseshoe

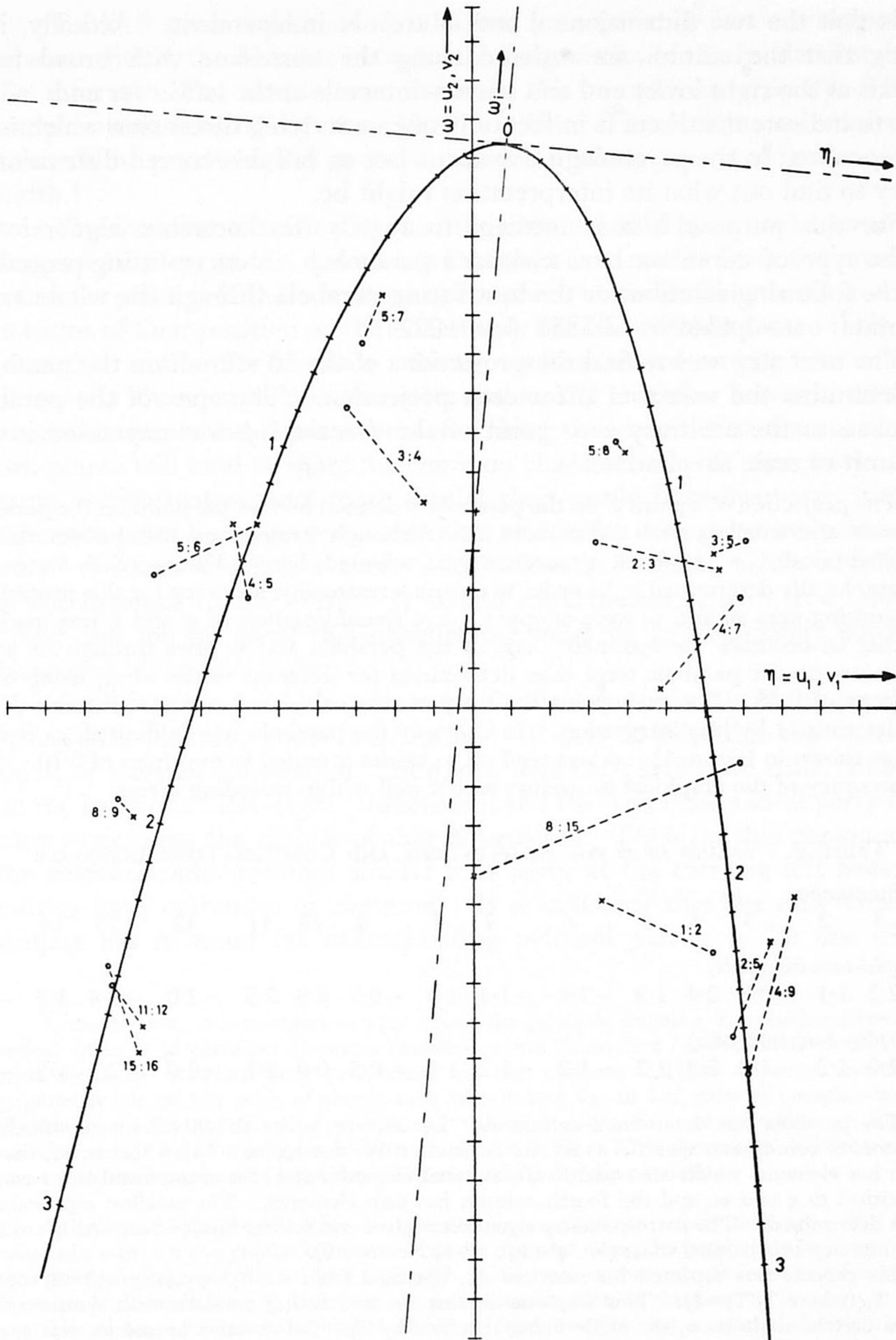


FIGURE 1. Superimposed projections of stimulus points on the joint plane $\eta = u_1, v_1$ and $\omega = u_2, v_2$. Corresponding simple-tone intervals (dots) and complex-tone intervals (crossed) are connected. The parabola $\omega_i = -3.3851(\eta_i)^2$ is given with working axes ω_i and η_i . Calibration marks on parabola indicate values of ψ in steps of 0.10.

S.P.

L

reveals that the two dimensions η and ω are not independent. Actually, it is striking that the stimuli are ordered along the horseshoe with broad tone-intervals at the right lower end and narrow intervals at the left lower end. This seems to indicate that there is in fact only one underlying dimension which is, as it were, curved in the psychological space. Let us call this curved dimension ψ , and try to find out what its interpretation might be.

For this purpose it is convenient to specify the horseshoe algebraically. For the type of curve we have chosen a parabola.¹ A curve-fitting procedure gave the following solution for the best fitting parabola through the whole set of 30 points: $\omega - 1.4346 = -3.3851 (\eta + 0.0374)^2$.

The next step was to find the projections of the 30 stimuli on the parabola, and determine the values of ψ for each projection. The apex of the parabola was taken as the arbitrary zero point of the ψ -scale. ψ was expressed in the same unit of scale as η and ω .

The projection of a point P on the parabola is defined here as the point on the parabola with the smallest real distance to P . Although projections can be determined algebraically, a graphical procedure was adopted here. Values of ψ were also graphically determined. In order to ensure a reasonable accuracy for this procedure, working axes ω_i and η_i were adopted, i.e. a transformation of ω and η was made so that ω_i becomes the symmetry axis of the parabola and η_i goes through its apex. Points on the parabola were then determined for different values of η_i , using equal steps of 0.05. For each point the tangent was calculated and the distance ψ was determined by line integration. In this way the parabola was calibrated, as it were (as shown in Figure 1). ψ was read off in values rounded to multiples of 0.10. The accuracy of the graphical procedure seems well within rounding errors.

TABLE 8. VALUES OF ψ FOR SIMPLE-TONE AND COMPLEX-TONE INTERVALS

<i>Stimulus number</i>		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
ψ_s (simple-tone intervals)		2.2	1.1	-0.8	2.4	1.2	-1.4	-1.4	1.3	-0.7	0.9	2.5	-2.0	-2.4	1.7	-2.4
ψ_c (complex-tone intervals)		2.0	1.2	-1.0	2.2	1.2	-1.2	-1.2	1.5	-0.5	0.9	2.1	-2.0	-2.5	1.9	-2.6

¹ The parabola was determined as follows. Let $\mathbf{A}_0 = (\eta, \omega)$ be the 30×2 matrix which has the vectors of coordinates η and ω as its two columns. We determine a 30×4 matrix \mathbf{B}_0 ; its first column has elements which are equal to the squared elements of η , the second and third column are identical to η and ω , and the fourth column has unit elements. The smallest eigenvalue of $\mathbf{B}_0' \mathbf{B}_0$ is determined. The corresponding eigenvector gives coefficients for the best-fitting parabola with symmetry axis parallel to ω : $(b_1 \cdot \eta^2 + b_2 \cdot \eta + b_3 \cdot \omega + b_4 = 0)$.

This process was repeated for matrices \mathbf{A}_i , obtained from \mathbf{A}_0 by applying a 2×2 rotation matrix \mathbf{T}_i (where $\mathbf{T}_i' \mathbf{T}_i = 1$). This way one obtains the best-fitting parabola with symmetry axis under a certain angle to ω , the angle being specified by \mathbf{T}_i . An iterative procedure was applied until the solution with minimum smallest eigenvalue was found. \mathbf{T}_i was varied in small steps corresponding to rotations of 0.001 degree, in the region from an angle of -30° to $+30^\circ$ with ω . Optimal solution was obtained for a rotation of approximately 5° (arc cos 0.084).

In fact, this parabola is not the best-fitting curve if one would allow also for other algebraic functions of second degree. The best-fitting second-degree curve actually is an ellipse. That the parabola was chosen in spite of this will be justified in the text.

Table 8 gives values of ψ . It was found that ψ has an almost perfect relation with d , the distance between the frequencies of the fundamental tones of the interval. For simple-tone intervals this correlation $r(\psi, d) = 0.970$; for complex-tone intervals it was $r(\psi, d) = 0.954$. The obvious conclusion is that intervals in psychological space are ordered along a curved scale of interval width.¹

The question remains why this dimension should be curved upwards. Our suggested answer to this question is that the curve results from the presence of a reference point. If there is such a norm point, stimuli will be compared not only in terms of their position on the linear scale but also with respect to their distance from the norm. For instance, take two points located on opposite sides of the norm. If there were no normative point, the distance between these points would be simply the distance measured along the scale. But if a norm exists the two points will tend to share the common characteristic of 'extremity' from the norm, which makes them more similar than would have been true without the reference point being present. In other words, for each point we have an 'extremity' measure which is a monotonic increasing function of the absolute value of the distance to the norm. It should nevertheless remain true that for two points not too far apart the interdistance must be approximately equal to that measured along the scale.

Such features would be accounted for by a curved scale in the psychological space, with a turning point at the norm.² It may be helpful in this connection to think of another illustration. Suppose political parties are differentiated only on the basis of a 'left-right' dimension, and that someone's ideal party is located somewhere near the middle of that dimension. Then for this person a party at the extreme right becomes similar to a party at the extreme left because both parties have extremity in common. It is still true that the only criterion the subject has in mind for differentiating political parties is the one dimension

¹ Another way to investigate to what extent the parabola explains the configurations in psychological space is to correlate distances between points themselves (as projected on the η, ω -plane) with distances between their projections on the parabola. These distances were measured graphically for all 105 pairs of simple-tone stimuli and for all 105 pairs of complex-tone stimuli. Correlations were calculated over squared values of distances. For simple-tone intervals this correlation is 0.969, for complex-tone intervals it is 0.908. This shows in a different way that the two configurations are to a large degree accounted for by the parabola.

² For a parabola $y = ax^2$, y can be taken to represent the extremity measure. If ψ is the scale value of points on the parabola, measured as the distance along the parabola to its apex (the normative point), we have $dy/d\psi = 1/[1 + 1/(4ay)]^{1/2}$. This means that rate of increase of extremity as a function of distance from the norm is an increasing function of extremity.

The choice of a parabola instead of some other function was a rational decision based upon the following considerations. First, we decided that an algebraic function of second degree would be the simplest solution. Secondly, although an ellipse was found to give a better mathematical fit, the parabola was preferred because we could not think of a meaningful interpretation of a closed scale. Closure of the scale somewhere opposite the normative point would imply that an extremely narrow interval would at the end become identical to an extremely wide interval; this does not seem to make sense.

'left-right', but the scale would be curved. In the limiting case we might even have a scale bent up so far that the two ends meet again, resulting in a closed loop.

One may argue that with a curved scale there is no longer a single dimension but two: extremity being a second independent dimension. This however is not true without qualification. The typical difference between an 'ordinary' two-dimensional space and the curved scale is that in the latter case points cannot be located anywhere in the space; possible positions are restricted to an area which follows the curve (allowing for error). If we have a curved scale the two-dimensional density distribution will have peculiar features in many respects different from, say, a bivariate normal distribution. In the tone-interval space it would be highly improbable to find a point somewhere midway between the two 'ends' of the parabola. Such a point would lie between narrow and wide intervals and at the same time be 'extreme'.

Finally it should be mentioned that in our study the existence of a curved scale was revealed by virtue of the existence of an objective measure for interval width. If such an external criterion is not available (e.g. political parties), differentiation between a curved scale and an ordinary two-dimensional space may well become difficult.

It can be therefore stated that what both interval configurations have in common is a plane on which the projections of stimuli are ordered along a parabolic curve in terms of interval width. The bending point of the curve, interpreted as being a normative reference point, corresponds with an interval width between the musical fourth and fifth, i.e., in the middle of the octave. Intervals which are not too far apart are differentiated in terms of interval width, but for larger distances it is necessary to make allowance for the 'extremity' of the interval (its distance from the normative point). Let us add that this interpretation finds some support in the fact that both very narrow and very wide intervals are unusual in musical compositions, so that 'extremity' might well be related to low frequency of occurrence. Indeed, the frequency distribution of musical intervals as they occur in three- and four-part compositions shows a mode around the middle of the octave.

(c) *Interpretation of specific dimensions* We found that u_3 and v_3 are uncorrelated dimensions. They are therefore specific to each configuration. For the complex tones it was possible, on the basis of earlier work on perception of musical consonance, to derive a hypothesis about the nature of the specific dimension. We shall discuss this hypothesis first.

(1) *The perception of ratio complexity* The reasoning starts from the old question of whether people are able to perceive the 'complexity' of a musical interval. Theories about consonance often state that the simplicity of a frequency ratio can be immediately perceived and that this would provide the basis for differentiating consonant and dissonant intervals.

In a previous article (Plomp and Levelt, 1965) it was shown that this simple ratio rule is valid only for complex intervals since it depends upon a

(presumably physiological) interaction between the harmonics of the fundamental tones. For simple tones the harmonics are absent; therefore simple-tone intervals are not differentiated on the basis of frequency ratio but only on the basis of frequency difference.

This leads to the prediction that in the configuration of complex-tone intervals ratio complexity would be present as a dimension. Moreover, it would be a specific dimension, since ratio complexity is irrelevant for simple-tone intervals.

To test this hypothesis we first had to define a measure of ratio complexity. Such a measure can be defined in several ways. For instance, one might take the larger value of the two numbers which form the ratio (after division by common factors). This measure was used in previous work (van de Geer, Levelt and Plomp, 1962), but in the light of what has been said above a more insightful measure can be developed.

If we have a simple ratio between the frequencies of the two tones in an interval, certain partials of these two tones will coincide. With a complex ratio, however, like 11:12, a larger number of the partials will be different and there will be a number of instances with small frequency differences between two partials. These small differences produce a kind of physiological interference which is thought to be basic to the simple ratio rule. That is, the ratio as such is not perceived; rather, there are auditory impressions which result from small differences between partials and these are more often present when the ratio becomes more complex.

With this in mind, a suitable measure for ratio complexity is the frequency of the first partial the two tones have in common. All further common partials will have frequencies which are multiples of that of the first one; our measure therefore is a measure of the 'wave length' of coincidences in the scale of the harmonics of the fundamentals.

The frequency of the first common partial, of course, is equal to the first common multiple of the two fundamental frequencies. Essentially being a product of two frequencies, this measure will have a skew distribution. Therefore, for our final measure of ratio complexity we took the logarithm of the frequency of the first common partial. This measure will be called a . In the following analysis it will often be easier to think in terms of *ratio simplicity* indicated by $-a$; this is simply the negative of a , or, if one prefers, the logarithm of the reciprocal of the frequency of the lowest common partial.

The measure of ratio simplicity having been defined, the next step is to find in both spatial configurations those dimensions which have maximum correlation with $-a$. This is simply a multiple regression problem. The solution is given in Table 9 where the entries specify direction cosines for the optimal dimensions, both for simple-tone and complex-tone interval space. These dimensions are called α_s and α_c , respectively. Table 9 also specifies the direction cosines for dimensions δ_s and δ_c which have optimum correlation with d , the measure of frequency difference.

The table suggests that in the configuration of complex-tone intervals α_c and δ_c mean different things: the angle between them is $\arccos(0.425)$. We also note that α_c is at a small angle to v_3 , the specific dimension in the complex-tone space.

TABLE 9. DIRECTION COSINES FOR DIMENSIONS WITH MAXIMUM CORRELATION WITH RATIO SIMPLICITY (α) OR WITH FREQUENCY DIFFERENCE (δ) FOR SIMPLE-TONE INTERVAL SPACE U (INDEX s) AND COMPLEX-TONE INTERVAL SPACE V (INDEX c)

	α_s	δ_s		α_c	δ_c
u_1	0.935	0.961	v_1	0.534	0.969
u_2	0.335	-0.253	v_2	0.191	-0.244
u_3	-0.119	-0.113	v_3	0.823	-0.032

For the configuration of simple-tone intervals, however, α_s and δ_s make a small angle ($\arccos(0.827)$), and α_s appears nearly perfectly orthogonal to the specific dimension u_3 . It is obvious, therefore, that the specific dimension in the complex-tone space is related to ratio simplicity, whereas this is not true for the simple-tone configuration. The fact, however, that ratio simplicity seems related to the common dimensions in the simple-tone space runs counter to the hypothesis which stated that ratio simplicity would be an irrelevant dimension for comparing simple-tone intervals.

Further qualifications have to be made when the correlations (Table 10) are examined. The projections of the complex-tone intervals on α_c have a high correlation with the values of a ($r(\alpha_c, -a) = 0.937$). On the other hand, the equivalent correlation for simple-tone intervals is much lower ($r(\alpha_s, -a) = 0.542$). The first correlation is in very good agreement with our hypothesis, but the latter is not: theoretically it should be zero.

TABLE 10. MATRIX OF CORRELATIONS

	u_1	u_2	u_3	v_1	v_2	v_3	$-a$	d
u_1		-0.049	-0.222	0.935	—	—	0.515	0.927
u_2			0.014	—	0.944	—	0.140	-0.259
u_3				—	—	0.064	-0.153	-0.270
v_1					0.099	0.132	0.610	0.891
v_2						0.073	0.271	-0.132
v_3							0.765	0.078
$-a$								0.534
d								

However, it can be shown that the relation between ratio simplicity and the simple-tone configuration is to a large degree an artifact. It happens that, for our selection of intervals, d and $-a$ are correlated; $r(d, -a) = 0.534$. Therefore, since simple-tone intervals can be differentiated according to their width,

ratio simplicity will be seemingly relevant for simple-tone intervals.¹ A test of this explanation is provided by correlating ratio simplicity and α_s after d has been partialled out. According to the hypothesis this partial correlation should vanish, whereas the equivalent partial correlation for complex-tone intervals should remain substantial. This is what is actually found. For the simple-tone configurations $r(\alpha_s, -a/d) = 0.215$; and for the complex-tone intervals

$$r(\alpha_c, -a/d) = 0.914.$$

The conclusion now seems clear. In the configuration of complex-tone intervals we have a specific dimension which is related to ratio simplicity (the correlation $r(v_3, -a) = 0.765$). For the simple-tone intervals the specific dimension is only slightly related to ratio simplicity ($r(u_3, -a) = -0.153$). In general, the observation that simple tones can be differentiated according to ratio simplicity is largely an artifact, due to the correlation between ratio simplicity and interval width.

(2) *The specific dimension for simple-tone intervals* Our final task is to find an interpretation for u_3 , the specific dimension in the simple-tone interval space. No hypothesis being available, the only approach open to us is an empirical one: to find a meaningful interpretation by trial and error.

The appropriate solution would seem to be the one indicated in Figure 2 in which values of u_3 are plotted against d . We find a W-shaped relation which is much too regular to be accidental. The interpretation of this curve, however, is

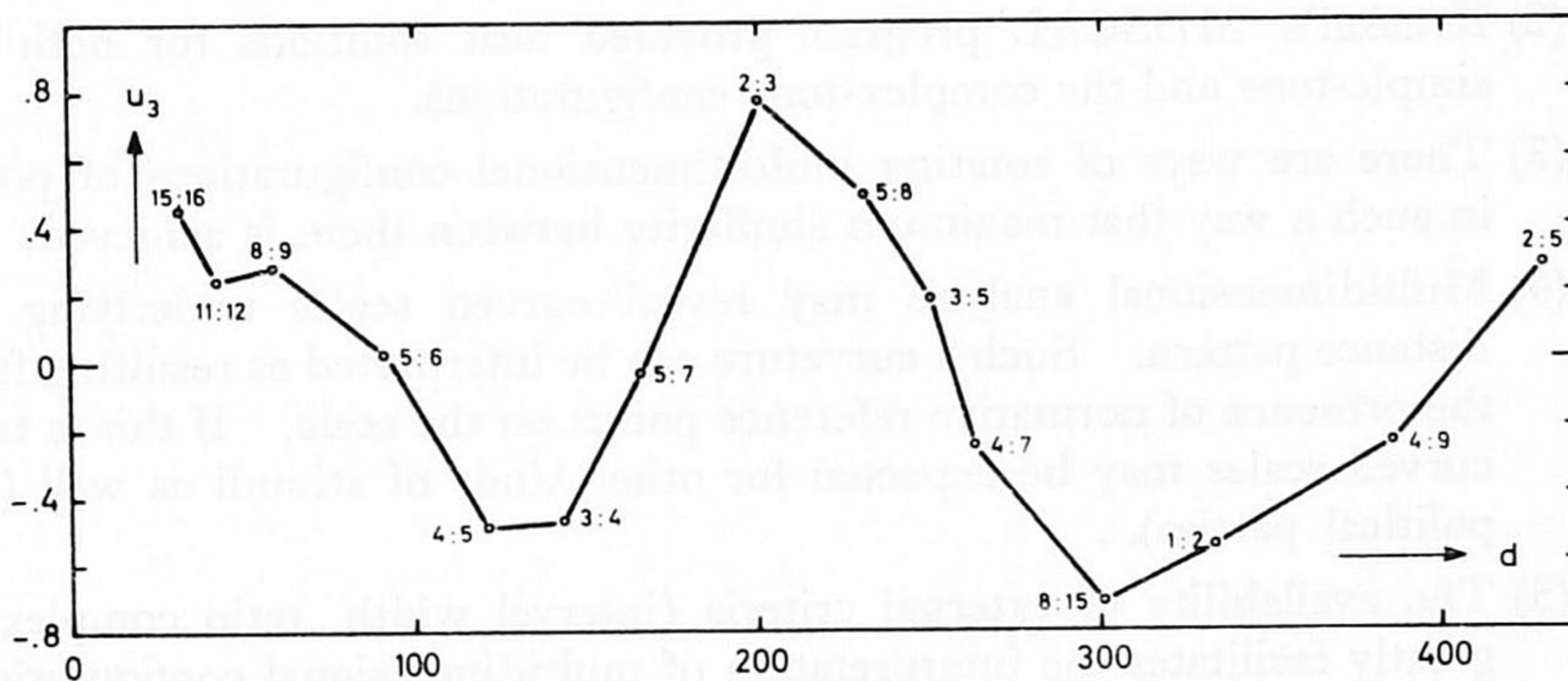


FIGURE 2. Specific dimension u_3 in simple-tone space as related to interval width.

not obvious. Our best conjecture at the moment is that the bends in the curve again result from the presence of reference points. According to this point of view the curve would reveal reference points around the octave, the fifth and the major third. Dimension u_3 differentiates intervals which have as it were, a

¹ This would be generally true. Narrow intervals will tend to have complex ratios; intervals with simple ratios will tend to be wide.

'third-like' character from those which are 'fifth-like', and higher up the scale differentiates those which are 'fifth-like' from those which are 'octave-like'. Why the curve remains in one plane is difficult to say. Perhaps this is due to the fact that an interval is never judged in terms of a 'third-octave' dimension because the fifth could be regarded as dividing the scale into two separate stretches.

That a curved scale of this kind should be specific to the simple-tone intervals is not hard to understand. Intervals near the octave can only have an 'octave-like' character for simple-tone intervals; with complex-tone intervals a 'mistuned' octave loses all resemblance to the octave because its structure of harmonics will be severely disturbed, and the same argument applies to the fifth and the third. However, these are only tentative conclusions which are far from being firmly established.

4. GENERAL CONCLUSIONS

This paper has a twofold objective. The first is an analysis of the perception of musical intervals, the second is a methodological one. As to the latter, a number of conclusions can be drawn.

- (1) For our Ss triadic comparison appeared to be an easy task. The incomplete design, divided over different subjects, was found to produce meaningful and consistent results.
- (2) Kruskal's MDSCAL program provided neat solutions for both the simple-tone and the complex-tone configurations.
- (3) There are ways of rotating multidimensional configurations of points in such a way that maximum similarity between them is achieved.
- (4) Multidimensional analysis may reveal curved scales underlying the distance pattern. Such a curvature can be interpreted as resulting from the presence of normative reference points on the scale. If this is true, curved scales may be expected for other kinds of stimuli as well (e.g. political parties).
- (5) The availability of external criteria (interval width, ratio complexity) greatly facilitates the interpretation of multidimensional configurations.

As to the perceptual aspects of our paper, the following conclusions are drawn.

- (1) Basically, musical intervals, whether consisting of simple or complex tones, are differentiated in terms of interval width. The apparent curvature of this scale is interpreted as resulting from the presence of a normative reference width which is possibly related to the frequency of occurrence of intervals in musical compositions (in the sense that both very narrow and very wide intervals are judged 'extreme' because they are unusual).

- (2) In addition, complex-tone intervals are differentiated on the basis of the complexity of the frequency ratio independent of interval width. This is not true (or much less true) for simple-tone intervals.
- (3) A further dimension in the simple-tone intervals space is tentatively interpreted as indicating the existence of certain ideal reference intervals: the major third, fifth, and octave.

REFERENCES

- BALL, W. W. R. (1939). *Mathematical Recreation and Essays*. Revised by H.S.M. Coxeter. London: MacMillan, 11th Edition.
- VAN DE GEER, J. P., LEVELT, W. J. M., and PLOMP, R. (1962). The connotation of musical intervals. *Acta Psychol.* **20**, 308-319.
- KRUSKAL, J. B., (1964 a). Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika* **29**, 1-27.
- KRUSKAL, J. B. (1964 b). Nonmetric multidimensional scaling: a numerical method. *Psychometrika* **29**, 115-129.
- PLOMP, R. and LEVELT, W. J. M. (1965). Tonal consonance and critical bandwidth. *J. acoust. Soc. Am.* **38**, 518-560.