U.S. NAVAL
MEDICAL RESEARCH LABORATORY

Submarine Base, New London, Conn.

Vol. XX, No. 7 REPORT NO. 353 27 April 1961

SCALING OF PITCH INTERVALS

by

J. Donald Harris

Bureau of Medicine and Surgery, Navy Department,
Research Project MR005.14-1001-2.07

SUMMARY PAGE

THE PROBLEM

To determine whether the mathematical scale used by musicians is truly in accord at all frequencies with the natural scale introduced by the human ear.

FINDINGS

The ear accepts the mathematical musical scale through the middle octaves, though not at the extremes; and this although discrepancies in the midranges are detectable by some ears.

APPLICATION

The results of this investigation have bearing on detection and discrimination abilities of the human ear in multitone situations.

ADMINISTRATIVE INFORMATION

This investigation was undertaken as a part of Bureau of Medicine and Surgery Task MR005.14-1001-2, Psychophysical Studies of Auditory Factors in Submarine Operation. The present report is No. 7 on this subtask. It was published in the Journal of the Acoustical Society of America, Vol. 32, No. 12, December 1960.

Reprinted from THE JOURNAL OF THE ACOUSTICAL SOCIETY OF AMERICA, Vol. 32, No. 12, 1575–1581, December, 1960

Scaling of Pitch Intervals

J. DONALD HARRIS
U. S. Naval Medical Research Laboratory, New London, Connecticut
(Received April 4, 1960)

A literature review showed that in bisecting a musical interval *Ss* tend to yield the arithmetic mean if the interval is large, but the geometric mean if the interval is small. Since fractionation judgments ("half-pitch") from which pitch scales may be derived would typically utilize quite wide intervals, and thereby tend to yield the arithmetic mean, a major discrepancy has arisen between such pitch scales and the pitch scale of our musical heritage, which is based upon the principle of the geometric mean. A number of experiments were performed using half-pitch and bisection judgments, and several variants of the method of equal-appearing intervals. From these it was concluded that when equal-appearing interval judgments are used with a standard interval no larger than about a musical third, a reliable psychological pitch scale emerges which agrees well with the common pitch scale of the piano keyboard. However, if the standard interval is as large as a musical fifth, the pitch scale begins to tend toward that derived from fractionation.

FOR at least two thousand years it was assumed that equal frequency ratios yield equal perceptions of pitch distance. Toward the end of the last century, however, certain psychophysicists called the fine details of the pitch interval sense in question, noting that nonmusical subjects often bisected a pitch interval by a frequency which was not the geometric mean (as demanded by our musical scale) but nearer the arithmetic mean. This last point, it was generally agreed, was a crucial one (Titchener[1]).

Unfortunately, Titchener did not see that in one of the papers he reviewed, that of Engel, the real clue to some experimental discrepancies was to be found. Engel found that the bisection of intervals smaller than an octave yielded geometric means, while with intervals larger than an octave the yield was progressively higher than the geometric mean.

Pratt[2] reported an extremely careful statistical determination. Each of three intervals was used: (1) an amusical interval, 300–410 cps, (2) 285–427.5 cps, and (3) an amusical interval, 285–510 cps. Pratt's results for intervals less than an octave were overwhelmingly in favor of the geometric mean.

To take account of interval-size Pratt[3] exactly repeated his earlier experiment but used four intervals over one octave but less than two octaves in extent. Three of his four *S*'s yielded bisections which fell progressively away from the geometric toward the arithmetic mean as the basic interval widened. The conclusion seems inescapable from these data that the bisection yield is a function of the basic interval size.

If, now, with intervals larger than an octave the bisection yield moves away from the geometric mean, then it must follow that a musical interval (a constant frequency ratio) will not have the same psychological magnitude as the same musical interval from another frequency region. Pratt[4] performed this validation experiment but with ambiguous results.

Although Pratt had all the concepts necessary to undertake the construction of a numerical pitch scale, it remained for Stevens[5] and his colleagues to invent the procedures and actually erect such an instrument. Stevens' first attempt consisted of having five *S*s set the frequency of an oscillator so as to sound at half the pitch of another oscillator. Ten standard frequencies from 125 to 12 000 cps were examined. These fractionations produced a pitch scale which differed considerably from the musical scale.

In a later paper Stevens[6] required interval quadrisectioning within the three regions 40–1000, 200–6500, and 300–12 000 cps. The results were transposed into a numerical pitch scale using the assumption that 20 cps is "zero pitch." This pitch scale, however, agreed only in the most general way with the fractionation scale of the first paper.

In an effort to resolve these differences Stevens repeated the first fractionation procedure using 12 *S*s (four were repeaters) and eight somewhat different standard pitches. However, a 40-cps tone was on tap at any time to help assess zero pitch. These data do indeed show a lowering in frequency of half-pitch judgments compared with the first attempt, and moreover furnish a fractionation pitch scale very close indeed to that erected by quadrisection.

A full-scale attempt to replicate and cross validate Stevens' 1940 mel scale was reported in abstract form only by Lewis.[7] He stated that trends or generality among scales were lacking: "A sort of 'generalized scale' based on many experiments involving different groups of subjects and different methods of observation may be found to have scientific utility; but such a scale will have to be used with appropriate caution."

In the present study we wished to discover (1) precise effects the several constant errors as enumerated by Stevens had on the mel scale, (2) the extents of individual differences and similarities for the several

[1] E. B. Titchener, *Experimental Psychology* (The Macmillan Company, New York, 1905), Vol. 2, Part 2, p. 241.
[2] C. C. Pratt, J. Exptl. Psychol. 6, 211 (1923).
[3] C. C. Pratt, J. Exptl. Psychol. 11, 17 (1928).
[4] C. C. Pratt, J. Exptl. Psychol. 11, 77 (1928).
[5] S. S. Stevens, J. Volkmann, and E. B. Newman, J. Acoust. Soc. Am. 8, 185 (1937).
[6] S. S. Stevens, *et al.*, Am. J. Psychol. 53, 329 (1940).
[7] D. Lewis, J. Acoust. Soc. Am. 14, 127 (A) (1942).

TABLE I. Half-pitch judgments. Entry: geometric mean of five judgments, standard frequency.

S	75	125	250	500	1 kc	2 kc	3 kc	5 kc	9 kc
CEG	56.0	98.5	156.3	242.2	465.4	1106	1786	2880	3733
KGW	56.1	82.2	101.9	192.9	302.5	949	1358	2126	3659
GNT	59.3	83.	122.7	140.1	198.0	233.	257	327	723
SES	62.9	76.0	161.8	320.5	606.3	1344.	1937	3230	5101
MHC	50.0	75.	125.	191.2	377.4	713	1202	2774	5102
CEW	57.5	84.6	166.7	332.7	661.7	1348	2079	3447	6158
EBH	53.4	67.4	120.9	176.5	274.1	406.	562	1594	3160
ARS	45.3	71.3	186.0	260.3	414.9	786.	1361	2075	3202
CKM	54.4	75.5	185.6	289.5	529.3	1313.	1866	2700	4468
JDH	54.7	81.1	172.3	300.2	491.9	1105	1858	2813	5125
Mn:	55.0	79.5	150	245	432	930	1427	2397	4043
Mdn:	56	80	160	250	437	1025	1575	2750	4000
Geom. Mn:	55	77	147	236	407	825	1231	2085	3601

procedures, and (3) the nature of the relations among the several available mel scales.

PRESENT EXPERIMENTS

Our first concern was with the difference between the Stevens' 1937 and 1940 mel scales. The factor of the added 40-cps tone to give an idea of "zero pitch" is complicated by individual differences known to be inordinately wide in judgments of this sort. It hardly seemed correct that the available low tone could be entirely responsible for the difference. We therefore arranged to have the same 10 *S*s make half-pitch judgments without any such tone, and again when it was forced upon them by the method of bisection.

Experiment I. Half-Pitch Judgments

A General Radio 913C oscillator was set to a standard frequency by a Stroboconn, while *S* adjusted the frequency drive of a Sonotone model AE21 audiometer to produce the variable frequency. Both tones were fed to a clickless (0.01 rise-fall time) electronic switch and timer which presented the standard for 1 sec, then after 1 sec silence presented the variable for 1 sec. Twelve seconds intervened before another pair was presented.

A loudness of 50 sones was maintained for all frequencies. The Sonotone frequency drive rotated a drum 4 in. in diam and 6 in. high mounted on top of the case, while the intensity drive operated a pen mechanism vertically on this paper. In this way, a Békésy-type audiogram could be drawn, or, after sufficient preliminary loudness matches, an isosonic contour could also be drawn. A 50-db isosonic contour was in fact drawn on the recording paper for each *S*. Then, as *S* slowly swept through any frequency range to make his judgment, the experimenter had only to move the intensity drive in such a way as to keep the dry pen riding on the isosonic contour. By this semiautomatic system we were assured that the many frequencies heard by *S* were all at 50-sone loudness.

*S*s were instructed to take plenty of time per judgment, and especially to be sure to bracket the frequency region between judgments "certainly more than half-pitch" to "certainly less than half-pitch."

The results for 10 *S*s are in Table I.

At first glance, the intersubject variability is seen to be extreme, so much so that it is difficult to choose the most informative average. We prefer the geometric mean.

Experiment II. Bisection Judgments

In this experiment, *S* adjusted tone *B* of an *A–B–C* series of which tone *A* was always a 50-cps tone 10 db over *S*s individual threshold. Standard frequencies 75 and 125 cps were not used since it was feared the 50 cps would not truly enough represent "zero pitch." In all other respects this bisection resembled the fractionation experiment.

The results for the same 10 *S*s are in Table II.

Figure 1 shows that neither the arithmetic means nor the geometric means indicate significant or even appreciable differences between fractionation vs bisection.

It is true that bisection vs fractionation yielded somewhat different curves for Stevens, as his Fig. 3 in the 1940 paper[6] shows, but if one examines the data for the same four *S*s who took part in the two experiments, such an overlap is exhibited that some appreciable if not major part of the difference in the Harvard experiments may perhaps have been due to individual patterns. In support of this argument, one notices that the frequency judged half of 150 cps in the 1940 experiment, where no 40-cps tone was used, was 85 cps, whereas that judged half of 125 cps in the former experiment was even higher, 90 cps.

Differences between our two procedures being statistically insignificant, the two sets of data were forthwith combined.

Table III presents the geometric means of all 10 judgments at each frequency for each *S*.

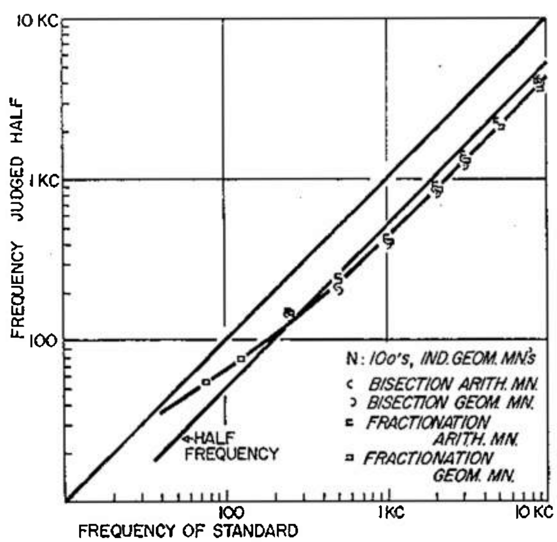

FIG. 1. Comparison between half-pitch judgments by bisection and by fractionation.

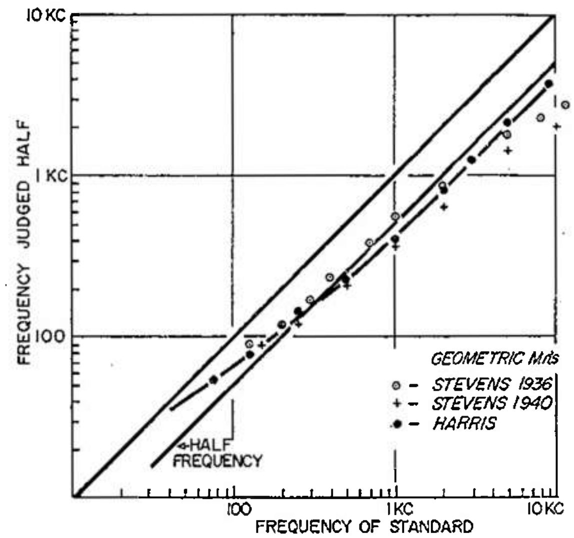

FIG. 2. Comparison of half-pitch judgments from three experiments.

Discussion of Experiments I and II

A comparison can now be made between these data and those of Stevens. When geometric means are computed and graphed for the three experiments, as is done in Fig. 2, it can be seen that our data interleaf the two previous sets except at the higher frequencies, 5 kc and above, where our *S*s yielded definitely higher frequency judgments. In fact, at 1 kc and above, our data are linearly related to log standard frequency.

There are other possible interpretations of these data, but the simplest is in terms of subject selection. Table III shows the rather extreme nature of individual differences. It would obviously be possible to extract almost any sort of mel scale depending upon what *S*s happened to volunteer or otherwise be made available.

But even with the lack of unanimity shown in Table III, certain features appear which seem to make a pitch scale possible as an average phenomenon if not as a prediction instrument for any individual. With respect to the half-frequency function these judgments are in large majority higher at 250 standard frequency and generally lower at 500 cps and up. Evidently something happens to pitch through the piano keyboard which yields half-pitch judgments now on one side, now the other, of the musical scale.

It seems to the writer, however, that the real question is not whether most *S*s can agree on a pitch scale but rather whether one can unearth and quantify the constants in the process of fractionation-bisection which in point of fact always do, on the average, yield such a pitch scale.

Acting on the suggestions in the work of Engel and of Pratt to the effect that much depends upon the width of the frequency regions to be compared, we planned a series of experiments to explore the method of equal-appearing intervals in pitch scaling.

TABLE II. Bisection judgments. Entry: geometric mean of five judgments, standard frequency.

S	250	500	1 kc	2 kc	3 kc	5 kc	9 kc
CEG	180.1	287.6	498	1126	2076	2814	5553
KGW	113.3	134.7	317.6	905	1492	2445	3611
GNT	129.9	173.3	236.9	244.	491.	604.8	1364
SES	132.5	152.1	506.4	1010	1504.	2693.	4565
MHC	113.8	158.8	259.1	761	1135	2254	4234
CEW	163.4	334.6	675	1370	2059	3368	6398
EBH	129.8	187.8	269.2	447	583	1810	3345
ARJ	156.6	201.8	402.8	789	1122	1554	2902
CKM	169.0	267.6	554.3	1146	1539	2581	4846
JDH	151.5	264.8	391.8	989	1155	2626	5064
Mn:	144.	216.	411.	879.	1316	2275	4188
Mdn:	142.5	195	400	950	1300	2350	4500
Geom. Mn:	142.	207.	388.	798.	1201	2095	3923

EXPERIMENTS ON EQUAL-APPEARING INTERVALS

Within this general method we distinguish two submethods to which we will give the self-explanatory names "adjacent extents" and "nonadjacent extents."

The equipment used in experiments I and II was elaborated to present a series of four tones in the usual 1-sec on-off sequence, with 8-sec frequency-adjustment interval between sequences. For this study, a Hewlett-Packard model 522B electronic counter was used to measure frequencies.

Appendix A shows why we abandoned the terminology and *aufgabe* of pitch ratios in favor of pitch extents.

Experiment III. Effect of Size of Standard Interval

It was first desired to know whether in fact the method of equal-appearing intervals was strongly dependent upon the size of the standard interval. If so, a glimpse would be had of certain discrepancies in the literature.

With one *S* three sessions were held with the method of nonadjacent extents (see Table IV for details), tone

TABLE III. Geometric means of all judgments of half-pitch. Entry: geometric mean of 10 judgments.

Subject	Standard frequency						
	250	500	1 kc	2 kc	3 kc	5 kc	9 kc
CEG	168.5	265.9	484.5	1120	1940	2867	4689
KGW	106.6	164.2	312.2	945	1435	2302	3702
GNT	127.4	165.1	241	243	490	680	1499
SES	147.6	252.2	556.4	1177	1721	2963	4833
MHC	119.4	175.0	318.2	737	1168	2514	4668
CEW	165.1	333.6	668.3	1359	2069	3407	6278
EBH	125.3	182.2	271.6	426	572	1702	3252
ARJ	171.3	231.0	408.8	788	1241	1814	3052
CKM	173.3	276.1	541.4	1237	1806	2715	4453
JDH	161.9	282.5	411.2	1047	1756	2719	5094
Mn:	146.6	232.8	421.3	908	1420	2368	4152
S.E.	7.8	18.8	43.3	114	174	248	418
Geom. Mn:	144.6	224.2	400.7	851.1	1290	2241	3908
Mdn:	154.7	241.6	410.0	996.0	1578	2614	4560

TABLE IV. Comparison of two methods with a large defining interval (1370 cents).[a]

Nonadjacent extents			Adjacent extents		
Interval	Freq.	Cents	Interval	Freq.	Cents
H	36		*H*	76	
		702			*625*
G	54		*G*	109	
		610			*517*
F	77		*F*	147	
		1011			*602*
E	138		*E*	205	
		806			*463*
D	220		*D*	268	
		937			*649*
C	378		*C*	390	
		786			*735*
A B	596		*A B*	596	
		1370			*1370*
Standard interval					
B A	1315		*B A*	1315	
		1587			*1798*
C	3318		*C*	3723	
		1301			*1183*
D	7042		*D*	7378	

[a] Subject ARJ. In collecting the data for nonadjacent extents, a session consisted of setting *C* such that *B* bisected *A–C*; then presenting *A–B–C* and setting *D* such that *C–D* = *A–B*. Here four tones are heard and *S* is instructed to ignore interval *B–C*. Then *A–B–D* is given and *E* set such that *D–E* = *A–B*; and so on. Five such settings of *C*, *D* . . . *I*, of which the geometric means are reproduced here. In collecting the data for adjacent extents the same five-setting geometric mean technique was used but *S* set *A–B–B–C*, *B–C–C–D*, *C–D–D–E*, etc.

A always being 893 cps. Three standard intervals of 1121, 1455, and 1973 cents were compared. In the downward direction, *S* set tone *C* (not the musical "C") to yield the following intervals: 893–734, 893–645, and 893–658 cps, respectively. But from a previous mel scale available for this *S*, it could be seen that these quite similar frequency regions had the very different values of 630, 855, and 1330 mels. These data thus show that pitch scales are indeed very sensitive to the size of the standard or defining interval, and that even the smallest of the intervals used here (1121 cents) was much too large for any one of the three mel scales resulting from this subexperiment to have any generality.

TABLE VI. Showing the effect on frequency settings *D* . . . *H* of relative settings of *C*.[a]

	C	*D*	*E*	*F*	*G*
5/28	435[a]	288[a]	181[a]	116[a]	55[b]
6/8	424	181	107	60	42[c]
5/29	422[b]	253[b]	178[b]	106	66[a]
6/2	369	255	185	94[b]	61
5/27	258[c]	153[c]	78[c]	40[c]	54

[a] Highest in column.
[b] Midscore.
[c] Lowest in column.

Experiment IV. Direct Comparisons Between Adjacent and Non-adjacent Extents

Subject ARJ was given both methods with a standard interval of 596–1315 cps (1370 cents) corresponding to no harmonic musical interval. Table IV gives convincing evidence that the two methods yield different results, and that the farther away one proceeds from the standard interval, the larger become the intervals by nonadjacent as compared with adjacent extents.

These conclusions were checked using two additional *S*s and three different nonmusical defining intervals. The results are in Table V. When comparisons are made between the two methods for all three *S*s at comparable intervals *C*, *D*, etc., it is seen that the superiority of adjacent extents is of the order of magnitude of three.

Experiment V. Optimum Generation of Tone C

A. Generation of an Average C by a Single-Session vs a Successive-Average Technique

It should be noted that, as would be expected, the setting of *D*, *E* . . . *I* is strongly dependent upon the original setting of *C*. The following Table VI gives the

TABLE V. Subject.

	ARJ				JDH				EBH			
	Nonadjacent		Adjacent		Nonadjacent		Adjacent		Nonadjacent		Adjacent	
Interval	Freq.	Av% dev. around the geom. mn.	Freq.	Av% dev. around the geom. mn.	Freq.	Av% dev.	Freq.	Av% dev.	Freq.	Av% dev.	Freq.	Av% dev.
G	54	15.4	109	4.4	141	20.7	260	.04				
F	77	36.0	147	7.6	216	8.6	324	4.3	68	13.7	175	4.5
E	138	32.2	205	2.7	323	5.4	402	1.4	103	24.9	243	1.6
D	220	22.1	268	4.6	454	10.7	480	1.8	244	13.2	351	9.3
C	379	14.4	390	3.4	672	2.4	616	8.6	529	7.8	560	3.2
A B	596				864				1000			
B A	1315				1436				2190			
C	3318	20.0	3723	24.9	1819	6.4	1897	5.0	3168	4.7	3293	4.0
D	7042	30.1	7378	14.0	2338	15.1	2368	7.1	4149	6.9	4332	5.9
E					2997	23.2	2844	7.5	6988	24.2	5992	10.1
F					3764	29.8	3685	3.3				
G					4757	35.8	4505	3.4				
H					5607	26.0	6036	3.1				

TABLE VII. Comparison of three variant methods of generating the interval B–C.

	Variant A (bisection) A–B–C^{a}		Variant B (adjacent intervals) A–B–B–C^{a}		Variant C (trisection) A–B–C^{a}–D^{a}	
	Lo–Hi	Hi–Lo	Lo–Hi	Hi–Lo	Lo–Hi	Hi–Lo
RHE	1181	623	1081	746	1092	746
JDH	1289	653	1125	685	1131	726
MHC	1222	668	1312	669	1312	669
CKM	1315	709	1134	761	1134	761
CEG	1081	757	1115	771	1105	760
JJO	1199	636	1078	737	1156	727
KGW	1319	534	1215	664	1273	664
BHC	*1146*	*649*	*1154*	*688*	*1182*	*655*
Arith. Mn:	1218	653.6	1152	715	1173	713.5

[a] Subject adjusts.

TABLE VIII. Method of adjacent extents, standard interval 878 cents. Entry: geometric mean of five independent trials.

Interval	JDH	CKM	SES	CEG	ARJ	EBH	Geom. mn.
G	260	249	244	252	300	136	224
F	324	301	312	333	342	194	296
E	402	378	416	386	403	271	372
D	480	479	526	480	496	402	476
C	616	595	670	626	615	576	616
A B	864						864
B A	1436						1436
C	1897	1817	2262	2429	2023	1909	2045
D	2368	2111	3123	3003	2537	2559	2594
E	2844	2380	4322	3732	3310	3244	3007
F	3685	2841	5492	4636	4076	3926	4027
G	4505	3392	8270	5550	5187	4966	5122

raw data from which column 1 of Table IV is derived. The protocols have been chronologically rearranged to show that once a relatively high C is set, its D . . . H will likewise be relatively high.

A direct comparison was made on one S as to the effects of generating C, D . . . H once in each session, as compared with generating five successive Cs and using the average for the generating of five successive Ds, and so on.

As accords with reason, the single-session procedure yields a much larger *variability* which tends to increase with distance from the standard interval. The variability in cents of the successive-average method, on the other hand, shows no trend with distance from the standard interval.

B. Three Variant Techniques for Generating Tone C

The question arises whether the precise details of how S generates the first subjectively-equal interval exert any great influence over successive intervals. Accordingly, JJO was given interval A–B and three variants were devised for him to generate the distance B–C.

In variant 1, S heard three tones A–B–C, and set the third such that A–B=B–C. In variant 2 S heard four tones and set the fourth A–B–B–C. In variant 3 S heard four tones but set both the third and fourth A–B–C–D. Table VII gives the values for C in cps.

Variant 1 yielded more variable data, and the average C moved further from B, as compared with variants 2 and 3 which were indistinguishable. In further work, variant 2 was always used.

Experiment VI. Pitch Scale by Adjacent Extents. Standard Interval 878 Cents

Six Ss were used to erect a pitch scale using adjacent extents according to the optimum variants explained in the foregoing. The standard interval was the smallest used previously, 864–1436, or 878 cents. (Appendix B shows why we started with a middle-frequency interval.)

Each frequency used to set D . . . G was the geometric mean of five previous trials, one per day. Results are in Table VIII.

The geometric means in the last column can be used to generate a mel scale which can be compared with any other mel scale, as we shall see later on. For our present purpose, the question of the final adequacy of experiment VI is whether the standard interval is too wide. For this reason we repeated experiment VI in essential details except that the standard interval was made 343 instead of 878 cents.

Experiment VII. Pitch Scale by Adjacent Extents. Standard Interval 343 cents

Using eight Ss, the method of adjacent intervals was expanded up to and including tone F, in which tone C was generated by variant B of experiment V. The results are in Table IX.

The data for individual Ss in Table IX have been

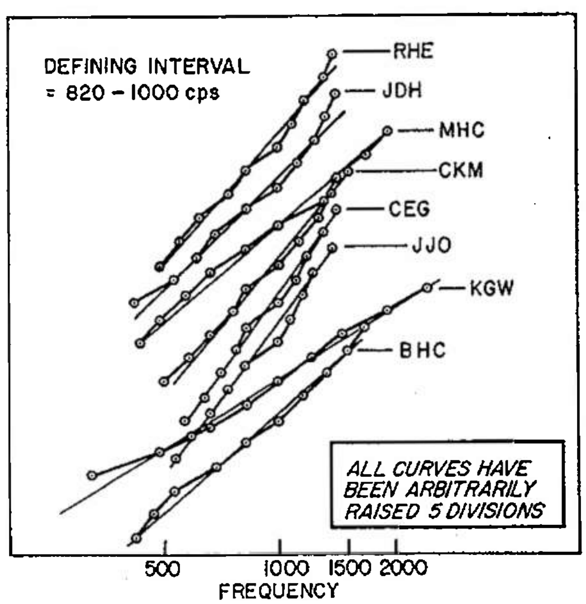

FIG. 3. Individual pitch scales using bisection with adjacent extent.

TABLE IX. Method of adjacent extents, standard interval 343 cents, C generated by variant B (see text). Entry: geometric mean of five independent trials.

Interval	RHE	JDH	MHC	CKM	CEG	JJO	KGW	BHC	Arith. mn.	Geom. mn.
F	497	422	440	506	570	540	246	425	456	444
E	552	527	496	584	646	598	328	472	527	517
D	625	616	578	666	705	664	490	533	612	606
C	746	685	669	761	771	737	664	688	715	714
A B	820									
B A	1000									
C	1081	1125	1312	1134	1115	1078	1215	1154	1152	1159
D	1189	1237	1483	1280	1196	1157	1456	1330	1291	1286
E	1319	1320	1700	1370	1308	1248	1929	1517	1465	1448
F	1399	1410	1950	1522	1420	1394	2448	1681	1653	1621

used to erect individual pitch scales in Fig. 3. The standard interval of 820–1000 was given an arbitrary value on the ordinate and the equal-appearing intervals have of course been assigned the same distance. The curves have each one been raised an identical arbitrary distance over the next lower one, in order to avoid crowding the graph.

Two facts stand out in Fig. 3, (1) that the mel scales are all to a first approximation linear with log frequency, and (2) that with one or two exceptions the slopes are very similar. The significance of (2) is that a method has been provided which yields between-subject variances of acceptable magnitudes far different than was found with fractionation-bisection as the reader has already seen in Table III. The significance of (1) is more fundamental in that, in contradistinction to all other pitch scales, it shows a frequency-pitch relationship exactly that of the musical scale.

With Table IX we are now in a position to compare graphically such a scale with others, and to argue the nature of discrepancies. The procedure involves the assumptions (1) that 30 cps represents "zero pitch" (though this is not critical—the scale will change but slightly with anything between 20–40), and (2) that the linear relation of pitch to log frequency can be extrapolated downward. Figure 4 contains such an extrapolation, with 1000 mels laid off linearly against log frequency between 1000–30 cps. With this mel scale, the geometric means of Table VIII were next converted into mels and also entered in Fig. 4, while the 1940 scale from Stevens has also been entered.

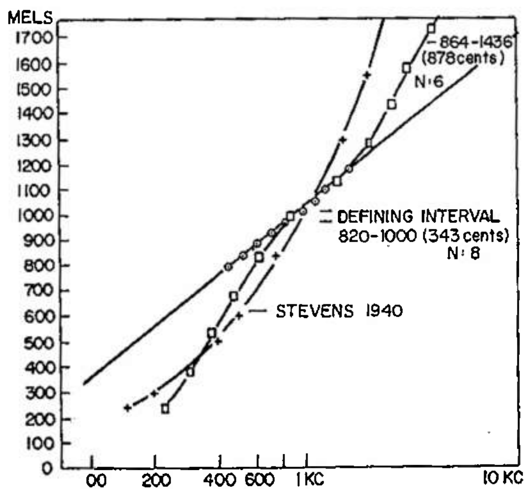

FIG. 4. Comparison of three mel scales.

DISCUSSION

A reasonable explanation now is apparent for the discrepancies among certain pitch scales: when intervals to be compared are in the same general frequency region (method of adjacent extents) and the standard interval does not exceed some critical value (of the order of magnitude of a musical third), pitch judgments are linear with log frequency as assumed by our musical confreres. When the standard interval exceeds this critical value, equal-appearing intervals at both ends of the standard interval will tend to cover a smaller frequency distance than predicted by the musical scale, and when the standard interval is extreme, as in bisection-fractionation, this constant error will produce a mel scale which differs by gross amounts from our common musical experience.

No doubt a musical literature could arise based upon a fractionation mel scale but it would have to consist of relatively large intervals—say, fifths (rather than tones and semitones) would be the smallest intervals utilized, while jumps of a twelfth would be common—and the melodic line would have to range in any one selection over several octaves more than now customary in most popular music. Furthermore, it would have to be monophonic unless a new harmony could be created. But under the best conditions it would sound, to most ears, worse than the bagpipes.

APPENDIX A. JUDGMENT OF PITCH RATIOS RATHER THAN PITCH EXTENTS

The question arises whether S could judge musical ratios more readily than musical extents. Such a procedure would have certain advantages.

A subject musically sophisticated, though not possessing absolute pitch, was utilized for a comparison of judgments of pitch extents vs judgments of pitch ratios. The same standard interval,

TABLE X.

Equal extents	Equal ratios
85	80
123	135
185	187
268	267
406	391
635	633
864	864
1436	1436
2089	1970
3379	3056
5083	4810
7560	7920

864–1436, was used in both cases. The S was given

$$\begin{matrix} 820\text{–}1000\text{–}1000\text{—}X\text{–} \\ A\cdots B\cdots C\cdots D\text{–} \end{matrix}$$

and asked in one *aufgabe* to set D such that $A-B=C-D$; the D was then presented as the C of the next set,

$$\begin{matrix} 820\text{–}1000\text{–}2089\text{—}X\text{–} \\ A\cdots B\cdots C\cdots D\text{–}; \end{matrix}$$

the new D was then presented as the next C,

$$\begin{matrix} 820\text{–}1000\text{–}3379\text{—}X\text{–} \\ A\cdots B\cdots C\cdots D\text{–}, \end{matrix}$$

and so on by nonadjacent extents. In the other *aufgabe* the same sets were given but subject was asked to set the fourth tone so that the ratio C/D appeared equal to A/B.

The results clearly indicate that these tasks produce identical results. For an average of five separate settings at each point, the data presented in Table X resulted with Subject GT.

No differences or trends appeared between the means or variances of these two sets of data. We conclude that it is not fruitful to seek for an especially valid pitch scale in the method of ratio judgment. On the other hand, it would probably be quite acceptable, if anyone wished to do so, to treat pitch-extent data as having been derived from ratio judgments and to use statistics appropriate to the latter in erecting a pitch scale.

APPENDIX B. JUDGMENTS PROCEEDING UP FROM ZERO PITCH

What turned out to be a false start was taken when Ss were presented with a standard interval A–B, in which A was set arbitrarily at 50 cps. This frequency was as low as could be generated cleanly in the earphone at anything like the purity and loudness desired, and seemed to be close enough to "zero pitch" as not to introduce a fatal error in the technique. Tone B was set at 350, 500, 600, and 700 in various sessions. Subject generated C by a form of bisection, and was then given A–B–C and asked to set D such that C–D=A–B. Then he was given A–B–D and asked to set E; and so on, equating the two end intervals and neglecting the third.

TABLE XI. Method of nonadjacent intervals, "zero pitch" = 50 cps. Entry: ascending intervals in cps, standard intervals.

	50–350		50–500		50–600		50–700
Interval	ARJ	EBH	ARJ	JDH	ARJ	KGW	ARJ
A	50	50	50	50	50	50	50
B	350	350	500	500	600	600	700
C	660	1087	916 872	864	1133	3667	2269 971 939
D	1053	1855	1303	1436	1844	4900	3536 1351
E	1276	2652	1802	2144	2466	5950	4800
F	1575	3555	2514	2874	3017	7250	7100
G		4164	3167	4250	3569	8100	8250
H		4300	3828	7300	4475		
I		4800	4650	9200	5100		
J		6100	5500		6100		
K		7000			7250		

Had this technique worked, it would have been possible to assign arbitrarily a value of, say, 500 mels to the frequency 500 cps, 1000 mels to the frequency C, 1500 mels to D, and so on, thus erecting a mel scale which had all the virtues of a numeral scale without the defects of fractionation.

Certain difficulties arose. In the first place, no S was willing to regard 50 as zero pitch. This might have been avoided by creating an acoustic system with which each S could generate what appeared to him to be zero pitch; but this was not done. In the second place, a comparison of A–B with, for example, F–G again introduces disparate frequency regions which was a main criticism of the half-pitch fractionation-bisection method. In the third place, the method reduced in no amount the wide individual differences which render the half-pitch judgments suspect. Some preliminary results on four Ss with the method are reproduced here in Table XI, each column being the results of one experimental session.

It is seen, even with these few data, that the between- and within-subject variances are as large as with any other method; and because of the indefiniteness of zero pitch, and because of the greater and greater frequency differences engendered, it was decided to abandon this variant in favor of one where the standard interval is in the central frequency region, the S being required to work both up and down from the standard interval.