# Training the absolute identification of pitch[1]

LOLA L. CUDDY[2]
*QUEEN'S UNIVERSITY, KINGSTON, ONTARIO*

*Two methods for training the absolute judgment of pitch, reference training and series training, were studied. Reference training concentrated during training on the identification of three reference tones in a set of nine pure tones, while series training gave equal weight during training to the identification of all nine tones. Results of pre- and posttraining tests, scored for the number of correct judgments, showed that reference training was more effective than series training for listeners with musical experience. In addition, discriminability (d') scaling of pre- and posttest performance indicated that reference training was particularly effective for training listeners with musical experience when the nine tones of a set were grouped into three pitch classes—high, medium, and low pitch. Listeners without musical experience benefited from both training methods, but their overall improvement was less than that for musical listeners.*

It has been well established experimentally that the absolute judgment of pitch improves with systematic training (e.g., Cuddy, 1968; Hartman, 1954; Vianello & Selby, 1968; Terman, 1965). The present study is concerned with the differential effects of two training methods—one method, reference training, concentrates during training on the identification of certain reference tones within a series of tones; the other method, series training, gives equal weight during training to the identification of all tones within the series.

The notion that development of a subjective reference standard tone assists pitch judgment of other tones has been suggested by both musicians and psychologists (Hindemith, 1946; Pollack, 1953; Seashore 1919) and has been tested experimentally (Cuddy, 1968, Experiment III). It was found that music students improved judgment of the pitch of 10 sine-wave tones if they were trained to recognize a single reference tone in the series. Series training in which each tone in the series occurred equally often, each tone was identified by the listener, and was immediately followed by knowledge of results, did not lead to improvement.

These results are surprising in that it is series training rather than reference training that preserves the

stimulus-response-reinforcement paradigm for associative learning. In the case of learning to judge pitch, however, the classical laws of habits may not apply, and models of structuring and organization must be considered (cf. Garner, 1962; Mandler, 1962, 1968). The success of reference training implies the presence of a cognitive tonal structure with the reference tone as a nodal point. For example, music students, who study the structure of Western tonality (that is, its intervals and scales), may develop a cognitive structure that maps the characteristics of musical structure.

The present study explores the generality of the effect of reference training on pitch judgment. Listeners were volunteers from university courses in psychology and were not preselected for musical experience. They were, however, classified as having some musical experience or interest (Group M) or having none at all (Group NM). The tones to be identified were determined according to their pitch distance in mels, but without reference to the musical scale. Half of the listeners in each group were subjected to reference training; the other half to series training.

Additional features were incorporated into the present study as follows:

(1) Stimulus-identification overlap, the correct identity of the stimulus being revealed immediately before or during presentation of the stimulus (cf. Aiken, 1967; Sidley et al, 1965; Swets et al, 1962) was incorporated into both reference and series training.

(2) The spacing of the tones along the stimulus continuum was experimentally varied. The question was, would identification of a set of tones, where the tones were grouped into three pitch classes, high, medium, and low, benefit more from reference training than identification of a set of tones spaced at equal pitch distances? Pollack (1952) has shown that the absolute judgment of pitch, which is normally limited to perfect identification of only about four tones, is not appreciably affected by the pitch spacing of the tones in the stimulus set. However, given the implied operation of structure in reference training, it was predicted that reference training would be particularly effective for learning to judge a set of tones in which classification of pitch was

emphasized by the grouping of tones along the frequency continuum.

(3) Empirical pitch scales were constructed for each listener. The scale values were obtained by a method analogous to the method for calculating d' from rating procedures (Green & Swets, 1966). Such scales have been developed for the absolute judgment of loudness by Durlach and his co-workers (e.g., Braida, 1968; Farrell, Pynn, & Braida, 1968) and are close to the theoretical notions of Creelman (1967). The purpose of the scaling procedure in the present experiment was to obtain a measure of the effects of training on discriminability between successive pairs of tones of a set where the measure was not confounded by response bias.

## METHOD

### Listeners

Listeners were 25 paid volunteers recruited from classes in psychology at Queen's University. There were 15 men and 11 women, ranging in age from 18 to 39.

Listeners were placed in Group M if they answered "Yes" (N = 13) and in Group NM if they answered "No" (N = 12) to the following question, "Have you engaged in any musical activities during the past 3 years?" Further questioning of listeners in Group M revealed that none was seriously studying music with the intent of pursuing a degree or a career in music. The modal response was, "I play the piano, have studied basics of music theory, and occasionally sing in choirs." Responses from Group NM included, "no success with music," "can't sing in tune," and "musical experience nil."

One listener in Group M was unable to complete the experiment, thus leaving 12 listeners in each group.

### Stimulus Tapes

Two separate sets of sine-wave tones were established, each set consisting of nine different tones. The tones of Set A were located within the range 400 to 2000 mels[3] and the difference between adjacent tones was 200 mels. The tones of Set B were partitioned into three groups within approximately the same frequency range, namely, 500 to 1900 mels; within a group of tones in Set B the difference between adjacent tones was 100 mels,

between groups, the difference was 400 mels. The frequency in Hz for the tones was determined from the revised pitch scale (Stevens & Volkman, 1940; Stevens and Galanter, 1957). Both sets were within the range 290 Hz to 3000 Hz.

Stimulus generation was similar to Cuddy (1968). The stimuli were produced by a Hewlett-Packard Audio Signal Generator Model 205 AG, and were recorded on Mylar Tape on a Roberts tape recorder Model 770. The frequency of tones in Hz was monitored during recording with a Hewlett-Packard frequency counter Model 3735 A. In order to reduce differential loudness cues, the voltage output of the signal generator was varied so that the SPL of the tones varied randomly over 15 dB. The tones were reproduced through monaural headphones at a comfortable listening level, averaging 70 dB SPL. The stimulus frequencies as reproduced by the tape recorder were checked with the frequency counter; it was found that the average deviation in frequency from trial to trial was within the limits of the accuracy of the signal generator and the frequency counter, ± 1 count at 1000 Hz.

Five test tapes constructed according to the method of absolute judgment (Garner & Hake, 1951) were recorded for each stimulus set. Each tape consisted of 81 stimulus presentations, nine tones each occurring nine times in random order. The duration of each tone was 1 sec, with an interstimulus interval of 4 sec. Training tapes each consisted of 54 double presentations. Each tone on a training tape was repeated, with 4 sec intervening between the first and the second presentation of the tone. Stimulus duration was held at 1 sec.

The tapes for series training consisted of the nine tones of a set occurring six times each in random order. Two series training tapes were constructed for each stimulus set. For the reference training tapes the tones in the second, fifth, and seventh ordinal position in each set (Tones L, M, and H) were chosen as reference tones, and were presented more often than the other tones of the set. Four levels of reference training, two tapes for each level, were recorded for each stimulus set. At Level I, the three reference tones occurred 14 times each, and the other six tones of the set occurred two times each. For Level II the corresponding frequencies of occurrence of reference tones and other tones were 12 and 3, respectively; for Level III, 10 and 4, and for Level IV, 8 and 5. Order of presentation of reference tones and other tones was randomly determined.

## Experimental Design and Procedure

There were four different training conditions, formed by the orthogonal combination of two training procedures, reference or series, and two sets of stimulus tones, Set A or Set B. Three listeners from Group M and three listeners from Group NM were assigned randomly to each of the four training conditions.

Listeners participated in six experimental sessions, over a period of about 2 weeks with no more than one session per day. The first (pretest) and last (posttest) session each consisted of three test tapes chosen at random from the five test tapes prepared for the set, or a total of 243 stimulus presentations. At the beginning of the session the nine tones of a set were presented to the listener in ascending order of pitch, and listeners were told that the names of the tones, in order, were L−, L, L+, M−, M, M+, H−, H, H+. Listeners were told that they would hear the tones played in random order, each tone occurring equally often, and that after each tone was presented, they were to record their judgment of its pitch in the response booklet provided. No knowledge of results was given during a test. The pretest session was preceded by a short run of 30 practice judgments to ensure that the listener was familiar with experimental procedure.

Sessions 2-5 consisted of training and testing, each listener receiving two training tapes followed by one test tape per session. For series training the temporal sequence of each double presentation of a tone was as follows: tone (1 sec)−listener's response (3 sec)−feedback (1 sec)−tone replayed (1 sec). Feedback for each tone was provided by means of a panel of lights in front of the listener. For reference training feedback was provided for reference tones only; listeners were required to respond to each tone by indicating whether or not it was a reference tone, and, if it was judged to be a reference tone, whether it was H, M, or L. All listeners started with the two tapes at Level I; if they made fewer than eight out of a possible 108 errors at Level I they proceeded at the next session to Level II and continued in this manner until they had completed the four training

sessions. The test tape following training was selected at random from the test tapes prepared, and listeners were required to name each tone without feedback.

## RESULTS

The results of the pretest and posttest for each listener were first scored according to (1) the number of correct judgments per tone, (2) the amount of information transmitted (T), and (3) the output uncertainty H(y). Sign tests for independent samples carried out on each of the three scores failed to reveal any differences in performance attributable to stimulus spacing, so for the following analyses the data were averaged across stimulus sets. The equivalence of stimulus sets is in line with the finding by Pollack (1952), but it was noted that the two highest scores of all the scores on the posttest, both in number correct, and T, were obtained by listeners in Group M given reference training on Set B.

The average number of correct judgments out of 27 for the pretest and posttest for Groups M (upper panels), and NM (lower panels) for reference (left-hand panels) and series training (right-hand panels) is plotted in Fig. 1 as a function of ordinal position of each tone. The average number of correct judgments for the entire pretest and posttest (possible total = 243) is given in Table 1, Columns 1 and 2. Inspection of Fig. 1 and Table 1 yields the following results, which are supported by an analysis of variance for repeated measures combined with a randomized block.

It can be seen from Table 1 that the difference between reference and series training is related to the amount of musical experience of group trained. Reference training is more effective than series training for Group M, but less effective than series training for Group NM. The interaction between musical experience, training method, and practice is significant at the .01 level [F (1,20) = 8.10]. Furthermore, the effect of training method on judgment appears to depend upon the ordinal position of the tone to be judged

Table 1

Average Number of Correct Judgments, Amount of Information Transmitted in Bits, and Output Uncertainty in Bits for Pretest and Posttest for Four Experimental Groups

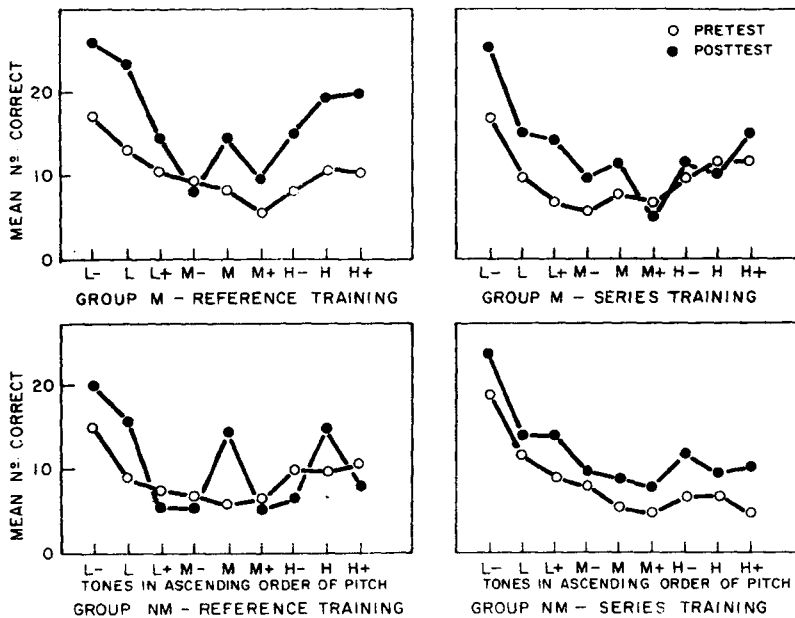| | Number of Correct Judgments | | Information Transmitted (Bits) | | Output Uncertainty (Bits) | |
|---|---|---|---|---|---|---|
| | Pretest | Posttest | Pretest | Posttest | Pretest | Posttest |
| Group M− Reference | 94.67 | 146.16 | 1.31 | 1.91 | 3.02 | 3.03 |
| Group M− Series | 90.50 | 122.33 | 1.26 | 1.62 | 3.06 | 3.11 |
| Group NM− Reference | 82.50 | 97.66 | 1.18 | 1.36 | 3.03 | 2.95 |
| Group NM− Series | 77.67 | 111.50 | 1.18 | 1.51 | 2.90 | 3.10 |

Fig. 1. Average number of correct judgments per tone for pretest and posttest for four groups of listeners.

mean $s_i$ and variance $\sigma^2$ on a subjective continuum. It was further assumed that in the absolute judgment task a response $R_j$ is selected by comparing the value of the random variable with each of a set of criteria (or response category boundaries) located along the continuum. The proportion of times a response $R_j$ or higher is assigned to a stimulus tone estimates the proportionate area of the distribution to the right of the category boundary for $R_j$ and may be converted to a normal deviate score to mark off normal deviate units along the continuum. Discriminability between tones $S_i$ and $S_{i+1}$ is then measured as the distance in normal deviate units between $s_i$ and $s_{i+1}$; this distance is independent of the actual location of category boundaries.[4]

The calculations for the discriminability scale were performed as follows: In order to compute T the data had been arranged in a 9 by 9 confusion matrix in which columns represented the nine stimulus tones and rows the nine response categories. Each entry in the matrix, $n_{ij}$, represented the number of times a particular stimulus tone, $S_i$, was assigned a

[F (8,160) Tone by Practice by Method = 2.37, p < .025]. The reference tones, H, M, and L were identified correctly more often in the posttest by listeners given reference training than by listeners given series training, and this finding holds for both Groups M and NM. The overall effect of practice, however, was greater for Group M than for Group NM [F(1,20) Musical Experience by Practice = 6.51, p < .025]. None of the other interactions reached significance at the .05 level. Significant main effects include the effects of practice [F(1,20) = 95.30, p < .001], the ordinal position of tones [F(8,160) = 21.80, p < .001], and musical experience [F(1,20) = 8.68, p < .01].

The average amount of information transmitted for pretest and posttest performance is shown for four experimental groups in Table 1, Columns 3 and 4. Results based on T values show the same pattern as results based on number of correct judgments. The rank-order correlation between T and number of correct judgments for all 24 listeners on the posttest was .91, significant beyond the .01 level.

Output uncertainty, or the average information in the listener's response, is shown in Table 1, Columns 5 and 6. Output uncertainty is less that 3.17 bits if the listener demonstrates a response bias. Inspection of Table 1 shows that output uncertainty increased as a result of practice only for groups given series training. This difference between training procedures is slight but significant [F (1,20) Method by Practice = 9.11, p < .01]. A clearer picture of this effect is shown in Fig. 2 where the mean relative frequency of each response

as a function of ordinal position of response is plotted for the pre- and posttest for all listeners given reference training (upper panel) and all listeners given series training (lower panel). From Fig. 2 it can be seen that for the pretest the function is bowed, with listeners tending to prefer the central responses. Following series training, listeners tended to distribute their responses more evenly across response categories except for a tendency to avoid the response "H+." Following reference training, however, there was a strong preference for those responses corresponding to reference tones, even though listeners had been told that each tone in the posttest would occur equally often.

The presence of response bias indicates that listeners given reference training adopted a strategy that was not optimal for maximizing the number of correct judgments when stimulus alternatives are equiprobable. The measure of information transmitted is also affected by response bias, since the value of H(y) places an upper bound to the value of T (see Garner & Hake, 1951).

## Discriminability Scaling

Discriminability scales were constructed to provide a measure of accuracy of recognition that was relatively independent of response bias. The method was derived from the theory of signal detection (Swets, Tanner, & Birdsall, 1961). It was assumed that corresponding to the presentation of a stimulus tone $S_i$, the internal event that determined the response could be represented by the value of a random variable that was normally distributed with
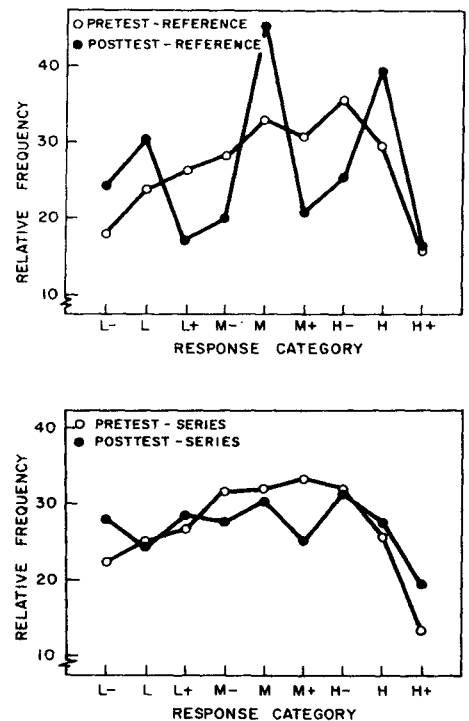


Fig. 2. Relative frequency of response categories for pretest and posttest for all listeners given reference training and all listeners given series training.

particular response, $R_j$. From the confusion matrix conditional probabilities $P(R_j/S_i)$ were computed; these values represented separately for each $S_i$ the proportion of total number of presentations of $S_i$ assigned to each response category. Cumulating the conditional probabilites across $R_j$ as $j = 9$ to 1 gave the proportion of times each stimulus was categorized above each successive response category boundary.

For each pair of adjacent stimulus tones, $S_i$ and $S_{i+1}$, the cumulative proportion above each category boundary for $S_i$ was treated as an inferred "false alarm" proportion, and the cumulative proportion above each category boundary for $S_{i+1}$ was treated as an inferred "hit" proportion. The corresponding hit and false alarm proportions for each category boundary were plotted on normal-normal coordinates, yielding an ROC curve with number of points equal to one less than the number of response categories assigned to both stimuli. A straight line could be visually fitted to most ROC curves so plotted, and a value of $d'(S_i, S_{i+1})$ was calculated as the absolute difference between the normalized false alarm and hit coordinates obtained at the intercept of the ROC curve with the negative diagonal. This procedure, in the case of unequal variances of the underlying distributions, gives equal weight to the units of each distribution (see Green & Swets, 1966, p. 98).

Values for $d'(S_i, S_i + 1)$ were obtained for each value of $i$ as $i = 1$ to 8, and the final scale value for each tone was obtained by cumulating the successive values of $d'$. The scale value thus represented (under the assumption of additivity) the distance of the mean of the distribution for each tone in normal deviate units from the mean corresponding to the lowest tone in the set. It may be pointed out that the scales derived from ROC analysis are logically equivalent to the equal-discriminability scales of Garner and Hake (1951) in that both procedures compute the distance between means of discriminable dispersions in standard scores. However, the Garner and Hake procedure obtains standard scores in terms of the average variance of all distributions that overlap at a given point on the subjective continuum while the present ROC analysis obtains standard scores in terms of variances of each successive pair of overlapping distributions.

Scales were constructed for the pretest and posttest for each listener, and scale values were then averaged across the three

listeners in each group for each training condition. Scales for listeners in Group M trained on Set A are presented in Fig. 3, left-hand panel. Scale value is plotted along the abscissa, pitch in mels along the ordinate. The pretest scales for listeners given reference and series training (open triangles and open circles, respectively) are quite similar. The function appears to be linear with the exception of a sharp downward deflection at the beginning of the scale (lower anchor effect). A slight deflection upwards may be noted at the end of the scale for the pretest of the group given reference training; this deflection was noted in the individual scales of two out of the three listeners in that group, but was not noted in the individual scales of any of the other listeners, nor for the posttests of any of the listeners given reference training. The functions for the posttest (closed triangles and circles for listeners given reference training and series training, respectively) may also be described as linear with a lower anchor effect. The posttest functions are of a steeper slope than the pretest functions, indicating greater discriminability between tones. The function for reference training is slightly and consistently higher than the function for series training, although the slope is not much steeper than the series training function.

The scales for listeners in Group M trained on Set B are presented in Fig. 3, right-hand panel. In constructing scales for the listeners given reference training scale values could not be determined for one listener in the group because no confusions existed between several pairs of tones. The functions in Fig. 3 for listeners given reference training are therefore the average values from two rather than three listeners. However, from the calculations that were possible from the data of the third listener, it was evident that her scale was of the same shape as the scales for the others in

the group, and the exclusion of this listeners from the group would not affect the conclusions to be drawn.

In Fig. 3 it can be seen that the functions for Set B are not linear. The two pretest functions for listeners given reference and series training show an initial sharp rise for the first three tones and a less steep slope for the remaining portion of the function. The posttest function for series training is higher than the pretest function, but the improvement seems generally due to increased discriminability between tones 600 and 700 mels, and tones 700 and 1100 mels. (Increased discriminability between tones 700 and 1100 mels is the only increase to occur in the individual scales of all three listeners given series training.) The posttest function for reference training, however, shows a marked increase in slope over the slope for the pretest function. The increase in discriminability among tones within the lower group, among tones within the higher group and between groups of tones is evident in the individual data for both listeners, and also for the third listener (not included in the plot) for whom only the upper half of the function, 1100 mels and upward, could be calculated. Discrimination among tones within the middle group is the only discrimination that does not appear to improve reliably in the individual data of all three listeners. Thus, it appears that reference training is particularly effective for training listeners in Group M on Set B.

Scales for Group NM were also examined; the results of these scales can be briefly summarized. All scales for Set A were of the same shape as the pretest scales for Set A for Group M; and scales for Set B were the same shape as pretest scales for Set B for Group M. Listeners in Group NM did not show an increase in slope from pretest to posttest. Improvement from pretest to posttest was due to an enhanced lower anchor effect,
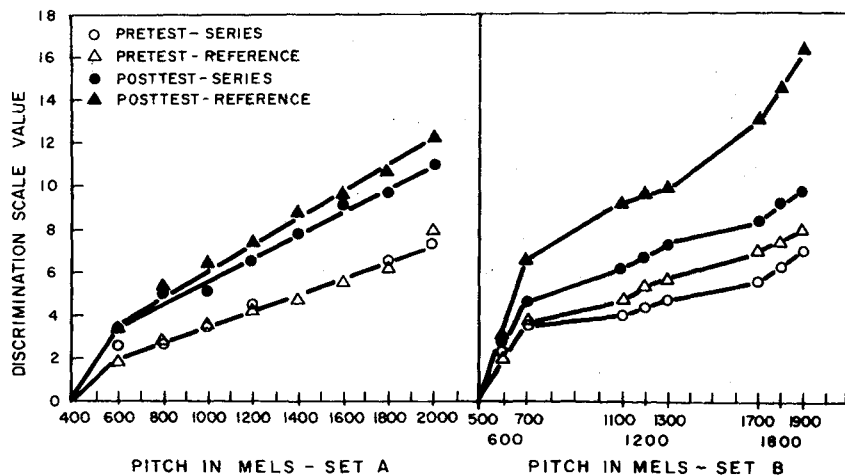


Fig. 3. Discriminability scales for listeners in Group M trained on Set A and Set B.

reflected in the posttest data by a sharper initial deflection of the scale. No difference between reference and series training was found. The benefit of reference over series training is peculiar to Group M listeners.

## DISCUSSION

The results suggest that in order for the judgment of pitch to be learned most rapidly there must be a processing system that has the capacity to perceive structure, a training method that initiates the adoption and development of structural rules, and a set of tones for which structure (i.e., interrelations of tones) is emphasized through the physical presence of musical structure or pitch classification as in Set B.

The structure of the pitch dimension may be specified in terms of Garner's definition of structure, "the totality of relationships between events [Garner, 1962, p. 141]." The lowest level of the system may be the channel capacity (Miller, 1956) of about three or four different pitches, or it may be a single reference tone. If the structure is a musical structure, the second level of organization might be the classification of tones within the octave, the third the classification of octaves. Such a system is reminiscent of the two-dimensional theory of pitch (Revesz, 1953; Shepard, 1964).

It was noted that in the present experiment series training resulted in improvement, whereas in the experiment reported earlier (Cuddy, 1968) series training was not effective. Because of the possibility that stimulus-identification overlap in the present experiment might account for the improvement, an experiment was conducted with randomly selected listeners in which three forms of series training were compared. In one training condition, knowledge of results (feedback) accompanied and overlapped the presentation of each stimulus tone; in another condition knowledge of results accompanied but followed the presentation of each tone; in a third condition knowledge of results was not given during training, but listeners scored their own responses at the end of each training session. Listeners in all groups improved their pitch judgment, but no difference was found among training conditions. It was concluded that stimulus-identification overlap was not critical for learning to judge pitch. Learning that occurs following series training is probably a very general effect due to increased familiarity with testing procedures. It is worth pointing out, however, that series training does not entirely prevent a listener from discovering a strategy of imposing structure on a set of tones. In yet other research in our laboratory we have twice noted listeners

who showed remarkable improvement under series training relative to the typical performance of listeners in the group. In both cases, these listeners reported without questioning by the E that they had tried to concentrate on one tone in the series, kept "imagining" the tone throughout the training and the tests, and used the tone as a reference to judge other tones. Both listeners were also trained musicians.

A final word may be added with respect to the discriminability scales. These scales are presented as empirical scales only. With respect to the assumption of decision theory, that the decision process is based on the outcome of a random variable $X$ whose distribution is Gaussian with standard deviation $\sigma$, inspection of the ROC curves indicated that the data were reasonably in line with the assumption. The ROC curves on normal-normal coordinates were straight lines of slope approximating unity. Slopes of greater than and less than unity were found, but the direction of the slope was not correlated with the physical frequencies of the stimulus tones. The shape of the scales is not an artifact of averaging of data; the scales for individual listeners were very similar in shape to the average scale. Finally, the assumption of additivity is problematic and remains to be tested. However, even if the assumption of additivity is not strictly correct, the main point is not affected: that for musical listeners discriminability between tones increases following training, and is particularly improved following reference training on a set of tones grouped into three pitch classes.

## REFERENCES

AIKEN, E. G. Prompting vs confirmation in the discrimination and identification of two auditory dimensions. Journal of the Acoustical Society of America, 1967, 42, 121-123.
BRAIDA, L. D. An equivalence between magnitude estimation and absolute identification for sound intensity. Journal of the Acoustical Society of America, 1968, 44, 384.
CREELMAN, C. D. Empirical detectability scales without the JND. Perceptual & Motor Skills, 1967, 24, 1079-1084.
CUDDY, LOLA L. Practice effects in the absolute judgment of pitch. Journal of the Acoustical Society of America, 1968, 43, 1069-1076.
FARRELL, R. M., PYNN, C. T., & BRAIDA, L. D. Determination of intensity discrimination and detection capabilities by absolute identification procedures. Journal of the Acoustical Society of America, 1968, 44, 389.
GARNER, W. R. Uncertainty and structure as psychological concepts. New York: Wiley, 1962.
GARNER, W. R., & HAKE, H. W. The amount of information in absolute judgments. Psychological Review, 1951, 58, 446-459.
GREEN, D. M., & SWETS, J. A. Signal detection theory and psychophysics. New York: Wiley, 1966.

HARTMAN, E. B. The influence of practice and pitch distance between tones on the absolute identification of pitch. American Journal of Psychology, 1954, 67, 1-14.
HINDEMITH, P. Elementary training for musicians. New York: Associated Music Publishers, 1946.
MANDLER, G. From association to structure. Psychological Review, 1962, 69, 415-427.
MANDLER, G. Association and organization: Facts, fancies, and theories. In T. R. Dixon and D. L. Horton (Eds.), Verbal behavior and general behavior theory. Englewood Cliff, N. J.: Prentice-Hall, 1968. Pp. 109-119.
MILLER, G. A. The magical number seven, plus or minus two. Psychological Review, 1956, 63, 81-97.
POLLACK, I. The information of elementary auditory displays. Journal of the Acoustical Society of America, 1952, 24, 745-749.
POLLACK, I. The information of elementary auditory displays, II. Journal of the Acoustical Society of America, 1953, 25, 765-769.
REVESZ, G. Introduction to the psychology of music. New York: Longmans Green, 1953.
SEASHORE, C. The psychology of musical talent. New York: Silver, Burdett, 1919.
SHEPARD, R. N. Circularity in judgments of relative pitch. Journal of the Acoustical Society of America, 1964, 36, 2346-2353.
SIDLEY, N. A., WINOGRAD, E., & BEDARF, E. W. Stimulus-identification overlap in learning to identify complex sounds. Journal of the Acoustical Society of America, 1965, 38, 11-13.
STEVENS, S. S., & GALANTER, E. Ratio scales and category scales. Journal of Experimental Psychology, 1957, 54, 377-411.
STEVENS, S. S., & VOLKMAN, J. The relation of pitch to frequency: A revised scale. American Journal of Psychology, 1940, 53, 329-353.
SWETS, J. A., MILLMAN, S. H., FLETCHER, W. E., & GREEN, D. M. Learning to identify nonverbal sounds. Journal of the Acoustical Society of America, 1962, 34, 928-935.
SWETS, J. A., TANNER, W. P., JR., & BIRDSALL, T. G. Decision processes in perception. Psychological Review, 1961, 68, 301-340.
TERMAN, MICHAEL. Improvement of absolute pitch naming. Psychonomic Science, 1965, 3, 243-244.
VIANELLO, MICHAEL A. B., & EVANS, SELBY H. Note on pitch discrimination learning. Perceptual & Motor Skills, 1968, 26, 576.