THE PERCEPTION OF PITCH IN MUSIC

by

Bruce Faulds

A DISSERTATION

PRESENTED TO THE

FACULTY OF PRINCETON UNIVERSITY

IN CANDIDACY FOR THE DEGREE

OF DOCTOR OF PHILOSOPHY

RECOMMENDED FOR ACCEPTANCE BY THE

DEPARTMENT OF PSYCHOLOGY

May, 1959

THE PERCEPTION OF PITCH IN MUSIC

Abstract

A battery of fourteen auditory tests was devised, partly from tests already in existence, partly from tests made up for this study. The tests examined various aspects of the discrimination of pitch, in particular in musical and nonmusical contexts. A fifteenth test was intended to measure the auditory digit span. The battery was given to two groups of subjects: (i) 67 freshmen from the Westminster Choir College, an institution specializing in music, (ii) 35 freshmen from Princeton University. The scores of the subjects on the separate tests were intercorrelated and factored. On rotation three factors identifiable as music, pitch, and memory were found. Factor scores were computed for the subjects and used as predictors of two criteria: (i) an arbitrary criterion which assigned a score of one to every subject in the first group, zero to every subject in the other. The attempt with such a criterion is to predict the institution of any given subject, or in other words to discriminate maximally the two groups. The multiple correlation was 0.68. (ii) The results of an eartraining test given at Westminster Choir College were made available to the author. A multiple correlation of 0.53 was found with this criterion. The main contribution to the testing of musical aptitude is in those tests which involve both pitch differences less than the keyboard intervals and a musical context.

ACKNOWLEDGMENTS

This study has been greatly facilitated by the help of many people. As his chief advisor for three years, including chairmanship of his thesis committee, the author would first wish to thank Professor Harold Gulliksen. Professors Carroll Pratt, Ledyard Tucker and Ernest Glen Wever also served on the committee, and to each of them the author is indebted for help and constructive criticism.

To certain faculty members of the Westminster Choir College the author's thanks are due. Dean Rhea Williamson showed an encouraging interest in the project: thereafter Mr David York assisted in many ways. He clarified the author's thinking during several cordial discussions, and scheduled the testing time with the freshman class. Other members of his staff in the Theory Department also assisted in providing grades.

Dr. Jack A. Vernon made available the freshmen in the Princeton sample.

Certain of the tests in this study have been used with the permission of their authors. In this connection the author wishes to thank:

Erik Franklin, Box 636, Askim, Sweden.

Robert W. Lundin, Hamilton College, Clinton, N.Y.

Harold G. Seashore, Psychological Corporation, 304 East 45th St., New York 17, N.Y., to whom the rights of the Seashore Measures now belong.

Herbert D. Wing, City of Sheffield Training College, Sheffield 10, and the

National Foundation for Educational Research, 79 Wimpole St., London W.1.

Finally to several friends and colleagues in Princeton the author is indebted for help in typing, making up tests, and administering them:

Douglas Candland proctored some of the testing sessions; and he and Mary Homrighausen played the piano for some of the recordings. Sara Matlack, Edna Bradbury, Ann King, Rosalee Hamilton, Geraldine Asbury and others were largely responsible for the typing and proofing of these pages.

The Printing Department of Educational Testing Service: James Conover, Joseph Nini and their assistants worked well and hard to meet deadlines.

Activities of this study were supported in part by Princeton University, the Office of Naval Research under contract Nonr-1858-(15), the National Science Foundation under grant G-642, and in part by the Educational Testing Service.

The study was performed while the author was a Psychometric Fellow of Educational Testing Service, 20 Nassau Street, Princeton, New Jersey.

THE PERCEPTION OF PITCH IN MUSIC

Table of Contents

THE PERCEPTION OF PITCH IN MUSIC

CHAPTER I

Introduction

In practically every field of knowledge the question arises as to whether it is better to run simple and straightforward studies, whose conclusions are unequivocal but not very much use, or to run studies whose potential usefulness is greater, but which also introduce difficulties because of their complexity. Thus in music sounds may be described in terms of duration, frequency, intensity and so on, and melodies in terms of rhythms and intervals; it is true also that these rather physical qualities do not suffice to explain the meaning and aesthetic appeal of music. A melody, for instance, has been defined (Schoen, 1925, p.36) as "... a succession of tones differing from each other in pitch and duration, and giving the effect of an aesthetic unity." Without the appeal to the aesthetic unity the statement would be clearly inadequate, and yet it is unsatisfactory as it stands, with aesthetic unity being to many people explanation by definition. In the psychological testing for musical aptitude both the simple and the sophisticated approaches are in evidence. Some investigators have examined the simple attributes of tone, such as pitch and intensity, knowing that repeatable observations could be had, and reasonably sure interpretations put upon them. Others have devised tests of a more sophisticated, musical, complex nature. Such tests have perhaps better face validity, but run the risk of having low reliability, and perhaps disputed interpretations.

This study takes one of the simple qualities of sound--pitch, which, if precedence can be given at all, is perhaps the most important single quality in music. It attempts to examine the sense of pitch in a variety of situations, some musical, others not, with the object of deciding the relations between these areas: whether people have one sense of pitch or several, how are they related, and how do we best predict success in music studies.

CHAPTER II

Review of the Literature

The best known battery of all is that of Seashore (Seashore, Lewis, and Saetveit, 1956). The tests in this battery are pitch, time, rhythm, timbre, intensity and tonal memory, and as such it examines what appear to be rather simple auditory qualities. The scores which the battery produces are reliable, but there has been much argument as to its validity for predicting success in the study of music. The field of music does not always produce as well-defined criteria as many other subjects: where this has obtained, the validity coefficients have rarely reached 0.40. Lundin (1953, p.208) gives a summary table of validity studies. Inspection of this table suggests that memory and pitch are the most valid subtests in the battery.

Kwalwasser and Dykema developed a somewhat similar set of tests in 1930. Several researches using the battery are reported in Kwalwasser, (1955).

Several tests which attempt to use the more musical material have since been developed. Some of them are no doubt improvements on Seashore, but others introduce difficulties of their own in that they depend on taste and learned distinctions rather than on objective and physically determined differences.

In 1920, one year after the Seashore tests Revesz (Franklin, 1956) had a rhythm test using musical material. It was given individually, and has been little used. He also used tests based on intervals, an idea that has been used by several later investigators. He also asked his subjects to analyze chords, and noted that many of the mistakes that were made showed that the subjects had been swayed by tonality: that is, they reported an analysis in accord with standard harmonic practice, even when a somewhat discordant stimulus had been presented.

König (1928) suggested memory tests analogous to the classical digit span test. These were apparently very discriminating, but were also given individually. He also found that tonal sequences were easier to repeat than atonal for both musically trained people and unmusical.

Schoen (1925) devised a test of tonal movement. Its items consist of four versions of a short melody. The beginnings are the same, but the endings are different, and the subjects are required to rank the four in order of merit. As such, this is one test in which the correct answer may be a matter of opinion. Franklin (1956) argues against the test on this score.

In a similar position are Lowery's tests (1926, 1929) of cadences and phrases. His tone memory test appears to assume too much sophistication in that the subjects are expected to ignore transposition, ornamentation, and tempo variations while still trying to observe more fundamental changes in the repetitions.

Wing (1948) has developed what seems to be the best tested battery after the Seashore and Kwalwasser-Dykema batteries. His tests are described as follows:

1. Detecting the number of notes played in a single chord.
2. Detecting and stating the direction of change of a single note in

a repeated chord.

3. Detecting changes of notes in a short melodic phrase.
4. Judging the more appropriate rhythmic accentuation in two versions of the same melody.
5. Judging the more appropriate of two harmonizations of the same melody.
6. Judging the more appropriate mode of varying loudness in two versions of the same melody.
7. Judging the more appropriate phrasing in two versions of the same piece.

They have been carefully standardized by ages in Britain. The last four appear to concern largely matters of taste; the third is similar in certain respects to Seashore's Tonal Memory; and the second will be discussed in more detail later in this report.

Drake (1933) added four tests to the field. Two of them treated of memory, a third of interval discrimination, and the fourth, which was called intuition, was not dissimilar to Schoen's test of tonal movement. Drake has argued strongly the importance of musical memory, and in 1954 (Farnsworth, 1958) the four tests were replaced with just two: one on memory, and one on rhythm.

Lundin (1949) has a series of five tests. The first of these uses interval discrimination, and is described in a later section; a second is based on the idea of melodic transposition, and as such is similar to Lowery's test. A third test called Mode Discrimination requires the listeners to note changes in mode in a repeated presentation of a chord, and as such seems to be less comprehensive than Wing's second test; and the last two tests depend on rhythm.

Franklin (1956) noted that one characteristic of Western music is its being written in a key. We commonly recognize a number of intervals as such; but in any simple tune these are probably superseded in importance by the relations of the various notes of the scale to the key-note, or tonic. It is usually possible, having heard the first few bars of a simple tune such as a hymn, to sing or play the final chord. The first few chords suggest to the listener an idea of the key, which then stays with the hearer through the tune. It is well known that one of the commonest techniques of the composer to introduce variety in his work is to change key. At all events it seems a reasonable assumption that those with little or no musical ability will have little key-sense, or in Franklin's terms, Tonal Musical Talent, while those with much will score correspondingly higher on a test of this Tonal Musical Talent, if such an instrument can be constructed. To measure this is the function of Franklin's test. It is conceivable that a sophisticated musician may get beyond the confines of tonality--one could hardly imagine a Stravinsky or Hindemith being bound by a key--and in this event it might not be too surprising if the scores on such a test were seen to fall at the highest reaches of musicality.

Factor-analytic studies. Several of the investigators mentioned above have gone on to correlate the scores of their various tests. Today the extension to factor analysis seems reasonable. As early as 1939 Drake (1939), using only Spearman's tetrad-difference criterion on a battery of eight tests found that a general factor and specifics were insufficient to explain the correlations. The tests involved pitch, time, intensity, rhythm and memory, and with such a variety of test material his result is not surprising.

Karlin (1941) reported the analysis of a battery of nine musical tests, mainly those of Seashore and Drake, and concluded that there were three factors. On rotation he defined these as tonal sensitivity, retentivity, and memory. This analysis would appear to have weaknesses: one of the tests, Emotional Sensitivity, is not described, but has its highest correlations with vocabulary and with Drake's Musical Memory. Along with the Seashore rhythm test and Drake's Retentivity it is used to define the retentivity factor. If we compute from Karlin's $\Lambda'\Lambda$ matrix the correlations of the primaries we get:

$$D(\Lambda'\Lambda)^{-1}D = \begin{array}{ccc} X & Y & Z \\ 1.000 & .538 & .265 \\ & 1.000 & .052 \\ & & 1.000 \end{array}$$

This suggests that two of the primary vectors have a considerable amount in common. We shall present confirmation of his factor of Tonal Sensitivity, which was defined by tests of pitch and interval discrimination. In our study the presence of certain other pitch tests has inclined us to call it simply Pitch. His memory factor is also paralleled in our study, and it is possible that the factor which Karlin called retentivity, and which is so closely correlated with the tonal sensitivity factor, was represented by tests involving both pitch and music. Drake's Retentivity test is certainly a rather complex instrument from its description (Drake, 1933). The addition of tests involving music in as pure a way as possible was one of the objects of this present study. Karlin also analyzed the table of correlations that Drake (1939) had computed, and tried to rotate to the same structure. He had no way of deciding whether the structure persisted, except from the appearance of the extended vector plots (Thurstone, 1947, ch. XI): in fact the retentivity factor of his own study is poorly represented, if at all.

In 1942 Karlin (1942) published a study of far broader scope, with tests intended to cover the whole field of auditory function. In addition to the commonly considered variables of pitch, loudness, quality and time he postulated domains of Auditory Analysis and Synthesis, Auditory and Visual Memory. The items of the tests in the Auditory Synthesis and Analysis domains used voices as stimuli, under various distorting conditions, the task being to interpret and write down. With a large battery he was able to define eight factors, including pitch, loudness, one composed of time and intensity and loudness together, three based on the analysis of speech and the nonauditory aspects of the battery generally, and two involving rather different aspects of memory. The tonal memory test from the Seashore battery was included, but this was, in fact, the most musical of the tests used. To this extent the study did little to clarify the retentivity factor of the previous musically oriented study: it did, however, establish the pitch factor.

In 1941 Wing (1941) analyzed his battery of seven tests by the simple summation method. This may be taken as giving approximately the same results as the centroid method or principal axes method of factoring. Following British tradition he did not rotate, but described his results in terms of a general factor accounting for 40% of the variance, plus two bipolar factors. The tests in his battery are all based on the piano. Three of them involve perception only, while the other four may require judgment of quality in addition. To see results comparable with U.S. traditions the author has rotated Wing's factor matrix to simple structure as follows:

	X	Y	Z
7. Judging best phrasing	.301	.000	.744
1. Detecting no. of notes in chord	.174	.436	.002
2. Detecting change of notes in chord	.423	.078	.249
3. Memory	.624	.002	.001
5. Judging best harmony	.000	.411	.260
6. Judging best loudness	.174	.077	.568
4. Judging best rhythm	.000	.066	.554

$D(\Lambda'\Lambda)^{-1}D$

	X	Y	Z
X	1.000	.667	-.106
Y		1.000	.213
Z			1.000

This would suggest a factor, X, involving memory, another, Y, involving harmony, the primary vectors of these two separated by only 48°, and then, more orthogonal, Z, a factor concerned with qualitative judgments. This conclusion is not greatly different to Wing's, but is perhaps more readily seen from the loadings. If we analyze the matrix by the principal components method, the roots in order are 2.8607; .9326; .2264; .1512; .1376; and there is a better case for 2 rather than 3 being taken as the rank of the matrix. This on rotation gives:

Factor matrix of rank two, from principal components factoring, of Wing's (1941) intercorrelation matrix. Rotated to simple structure.

Test:		
7.	.092	.690
1.	.665	.010
2.	.537	.137
3.	.818	-.201
5.	.294	.350
6.	.040	.585
4.	-.148	.574

$$\Lambda = \begin{vmatrix} .550 & .453 \\ -.835 & .891 \end{vmatrix}$$

The distinction here again seems to be whether the tests call for judgments or perceptions. The angle between the primary vectors is about 60°.

The study in which Franklin (1956) describes his test of Tonal Musical Talent is also factor analytic. Two studies were run. In the first, of thirteen variables, four factors were extracted. One of them is probably the same as the judging factor or function of Wing's tests--these being included in the battery--another loads highly on rhythm tests, but the last two are of most interest, in that the first hinges on pitch, and the second on memory.

From the second correlation matrix of eighteen variables eight factors were extracted. This number is undoubtedly high, but the battery involved a rather large variety of material, including four intelligence-type tests. The patterns of loadings are not as clear as in the first study, but the area of pitch-music-memory is at all events pretty much restricted to three of the eight factors--A, C, and G in

Franklin's listing.

McLeish (1950) in a comprehensive article factor analyzed the Seashore battery. In his first factor pitch and memory emerged as the subtests loading highest with more musical material, such as the composite score from Wing's battery. He also compared the contributions to the variance of the factors when correlation matrices of Wing's and Seashore's tests were factored in the same way, concluding that the Wing battery was more homogeneous, as indicated by its large general factor, while the Seashore had a smaller general factor, but larger specifics. This contention is reasonable if we bear in mind that all of Wing's battery is based on the piano, and a goodly part of it on qualitative judgments: while the Seashore battery has for long been known to consist of rather specific tests.

McLeish also examined the effect of age and intelligence on the results of music test scores by computing the correlation table with these partialled out, concluding that they had little effect.

Many years before, Brown (1928) had also used partial correlation as a means of determining that age or intelligence have little influence on scores on Seashore tests. Brown's report was not factor analytic, but he did find that with a group of 105 high school students, the Memory score correlated highest with the teachers' ratings.

Growing out of the research set in motion by the war Harris (1957) has reported studies on audition. In one 27-test study he found seven factors. Pitch was one of these, though it had a strong memory component: and the factor of next most interest was called melodic memory.

On this factor tonal memory tests have large loadings. The interrelation of pitch and memory is described by Harris: "We knew that pitch was a factor, and believed it (erroneously) to be a single unitary trait. We knew that memory for tone progressions was distinct from pitch memory or pitch discrimination, and from other types of auditory memory." (Harris, 1957, p. 6).

König (1957) has reported comparative results of several methods of determining the differential limen for pitch. He concluded that if the Seashore paradigm is used, a time interval of several seconds can be introduced between the first and second stimuli before any subjects' scores change notably, and that as this interval increases, some subjects' scores fall more than others'. This suggests the operation of two factors, one based on memory, and the other on pitch.

About the same conclusion was reached by Bachem (1954), though his subjects were musical--half indeed had absolute pitch. Bachem was accordingly able to have much longer time delays than Konig, but found that in the range of musical frequencies the most musical subjects had better memories for pitch.

This survey has concentrated on those studies and tests which seemed most related to the way in which music is heard, with particular reference to pitch. It has ignored experimental work on aesthetics (Farnsworth, 1958), tests of music knowledge (Lundin, 1953), and tests of written music (Aliferis and Stecklein, 1955). All of these areas are concerned with how we perceive or understand, rahter than hear, music.

CHAPTER III

The Problem

The examples noted in the literature of investigations of academic knowledge and aesthetic appreciation are doubtless of importance in assessing the individual's enjoyment or ability, but bypass an earlier question which does not yet appear to have been satisfactorily answered. The question concerns the mechanisms by which music and sounds are heard. The computer analogy or information channel is a popular one today. In these machines the input is of a certain form, punched cards, punched tape, magnetic tape; it is read into the machine by various devices--sensing brushes, magnetic read heads--and is then processed by the machine's circuitry: in these terms we are interested in the way that the machine reads its input, rather than the input itself, or the subsequent processing of the input data.

Several of the writers mentioned above have speculated as to the mechanism by which the hearer might perceive the nature and meaning of the sound waves impinging on his ear. Seashore's opinion was that people possessed a number of specific abilities, largely inherited, and not greatly subject to training. The subtests in the battery then measured the extent to which the individual possessed each of the abilities. Since that time it has generally been conceded that training does have its effects--even in the disputed area of absolute pitch (Neu, 1947), but the first two hypotheses must still be regarded as possibilities. The attacks on Seashore on account of the atomistic nature of his tests have not shown him to be wrong, though they have brought to light other writers' theories.

Some of the analyses quoted in the previous section have suggested a larger or smaller number of factors composing our sense of music. These factors were not as simple as those that underlie Seashore's tests.

At the other extreme, having noted the low correlations with grades that music tests often show, Farnsworth once tried (1935) to decide whether intelligence was not as much use as music tests in the predictions of various types of music grades. In the study he used only two of the Seashore tests, pitch and memory, which is perhaps a restricted sampling of music tests, but these are very probably the two best in the battery for the purpose.

Another standpoint is taken by Revesz, (1954). To him the musical person was characterized by the mental conquest of music as art. This meant that music was not passively assimilated, but that along with the hearing of it there had to be a heightened mental activity, coordinating, analyzing, seeing structure and style, and fashioning one's being to the mode of the music. This description is expanded in Revesz's book, but enough has probably been said for the general feeling to be apparent. He could conceive of unmusical people who might have very fine pitch discrimination and so forth. His several tests accordingly fell into the more musical type, with relative pitch, chord analysis, etc. He was of the opinion that there was but a single talent, and that all of the tests were an attempt to assess this single talent, though he made no use of correlational and factor-analytic techniques to test this assumption.

The weight of the evidence so far suggests:

(1) that the auditory domain is on the whole very complex. Seashore

originally assembled his battery so that the mutual correlations of the subtests would be low, and many other types of tests since added to the field have also shown themselves to be rather specific.
(2) that of the Seashore battery pitch and tonal memory are the best predictors of musical ability.
(3) that among the more recent tests there has been on the whole an insufficient attempt at direct validation, such as multiple regression on a criterion. Reports like Franklin's and McLeish's are in some sense validation, in so far as they compare Wing's and other tests with the Seashore in factor structures. One of the difficulties of validating a test by means of an independent criterion in that the criterion itself might be in error. Such criteria are often enough teachers' grades, or based on tests which might, in fact, have little to do with the ability they are assumed to test. It seems probable, for instance, that the Farnsworth study just quoted utilized poor criteria.
(4) if we take the Wing battery to be an example of the more musical tests, then the McLeish report suggests, in fact, that there may not be as much difference between the atomistic and musical tests as Seashore's critics have suggested.

For these reasons a study was planned that would:
(1) throw light upon the manner in which a series of tests, partly musical in content, partly physical, would work together. It would pay particular attention to the borderline between music and pitch, with tests both musical and involving small frequency differences. This borderline has not, to the author's knowledge, been previously investigated in a factor analytic study. Of the studies mentioned

in the Review of the Literature, particularly those of Karlin (1941) and Franklin (1956) both musical and nonmusical tests are used, but the pitch factor is tested for by tests involving only pitch, and similarly for the music factor. What sort of factor structure would be produced by such a battery, and which tests would be the best predictors of musical achievement?

(2) pay attention to the role of memory and of pitch, and of their interaction.

(3) be able, hopefully, to answer questions of validity by examining subjects whose musical ability could be independently judged, and would show wide variation.

CHAPTER IV

Design

When we have a few concepts which might be independent, or might be so closely related as to be two names for the same thing, factor analysis is a design to be considered. This is particularly the case when the concepts cannot be exactly described, but--until more information is available--represent a conjectural and somewhat poorly defined quantity. Even with a less than exact definition of one of these quantities it may still be possible to decide that the conjectured ability will enter into a number of other, known, measures. The pattern of correlations of all the measures can then be inspected to see two things. Firstly: those measures which were put in to stand for one of the hypothesized domains--how closely do they in fact stand in relation to each other? If the members of such a group of measures have very similar factor loadings or correlations, this is evidence that they do all measure a single underlying ability. Secondly: if these primary domains are identifiable, how nearly are they related to the other domains? This would be described in factor analytic terms as the angles between the primary trait vectors, in a Euclidean space of dimensionality usually equal to the number of primary concepts that have been hypothesized. Often in addition it is possible to devise tests which will contribute to more than one domain, and if these prove to have been soundly conceived and constructed, the experimenter's position is strengthened by the appearance of a complete and convincing simple structure.

Just at this point we run into a difficulty: if we wish to compare musical with nonmusical tests, and the role of memory in both, it would be ideal to define the corners of the hypothesized configuration with the purest possible tests of music, memory, and so on, that we can devise. However music with the variable of pitch held constant ceases, to all intents and purposes, to be music. In this artificial case rhythmic patterns and tempos may persist, and changes of timbre on a given note are certainly possible, but the most important thing has been lost.

Similarly about the only case in which two stimuli can be presented to the ear without invoking memory are the dichotic experiments (Wallach, Newman and Rosenzweig, 1949), which introduce complications of their own. This means that a certain amount of compromise may be necessary in the selection of tests.

Other considerations which influenced the selection of tests were testing time available, and the desire to use group testing sessions with multiple choice pencil-and-paper-type items.

The test battery

Description of tests

Some of the tests in the battery are described in the literature by their authors, and for these the descriptions are accordingly brief. Also in the Appendix a complete test booklet is given, and reference may be made to this for details of test directions.

A listing of the tests is given with short descriptive names: this is followed by more complete descriptions in terms of the hypothesized factor structure.

1. Pitch test, from Seashore
2. Detection of changes in a chord, from Wing
3. Interval discrimination, from Lundin
4. Detecting departures from true octaves
5. Detecting departures from true scales
6. Detecting departures from the true, of a tune
7. Pitch test, masked by white noise
8. Tonality, from Franklin
9. Pitch test with bands of white noise
10. Pitch test with time delays
11. Memory in chord sequences
12. Timbre test
13. Tonal memory, from Seashore
14. Timbre, with time delay
15. Auditory digit span.

Pitch, divorced as far as possible from Music and from memory, was represented in the following tests:

1. A shortened from of the Seashore pitch test. A 78 rpm record was taped, and 35 items were taken as the test. Numbers were read aloud each five items.

9. Bands of white noise. This idea was used in a more crude form in the Karlin (1942) study. He had a number of noises of various descriptions on record, and the subjects had to say which of a pair sounded on the whole higher. Karlin used the majority decision to key the items. The advances in electrical apparatus make it possible today to produce white noises of any desired spectrum, and this avoids the necessity of keying items on subjects' judgments. Such a test was developed during the war (Karlin & Stevens, 1946), but it was no longer available, and the present test was made up specifically for the study. A white noise generator was connected to high and low pass filters and lengths of tape of different noise bands were recorded. From these the items were spliced together. The choice of cutoff frequencies was arbitrary, except that a range of items was chosen that would hopefully be discriminating as regards difficulty, and which would not use frequency variations of such a size that allowance would have to made for intensity. In fact an easy item was formed by the bands

$$\left.\begin{matrix}1200\\900\end{matrix}\right\} c/s \qquad \text{and} \qquad \left.\begin{matrix}1300\\1000\end{matrix}\right\} c/s$$

and a difficult item by the noise bands

$$\left.\begin{matrix}1200\\900\end{matrix}\right\} c/s \qquad \text{and} \qquad \left.\begin{matrix}1220\\920\end{matrix}\right\} c/s$$

One or two observers were asked at this stage to judge on intensity as a check, but could detect no differences.

7. This test was similar to test 1, except that a background of white noise was superimposed. It was originally included in the battery to see whether the tests would group into two sections depending on whether or not any irrelevant stimuli were competing for the ears' attention. This particular split is not evident in the structure, and is for that reason little considered further in the rest of this report. In any event the choice of items for this particular test proved to have been too easy, to the extent that the test was relatively worthless.

Music cannot be divorced from pitch, as we have seen, but the structure of western music is such that only a certain number of pitches is used. On many instruments such as the piano the intermediate pitches cannot be produced at all, and on instruments such as the violin where it is possible to produce them, to do so is condemned as a fault. However the semitone, which is the smallest interval on the keyboard, is, in the middle ranges at least, within almost everyone's powers of discrimination. It seems a possibility then, to devise tests that will use only these intervals, and will therefore not depend on the subjects' abilities to discriminate fine shades of pitch. What they are discriminating in this case is open to question. However let us now describe one or two tests which it was hoped would define a corner of the configuration, to be associated with a Music factor.

8. This is Franklin's test of Tonal Musical Talent. In this test the items consist of the beginnings of melodies. The subjects are asked to consider their endings and say whether they are wrong, right, or left out. The manner in which the items are built up makes the rationale clear. In the case of a <u>right</u> ending item the first part of the item will establish the key, and the final chord will be the tonic just

established. Similarly in a wrong item the last chord played will be discordant, or will at all events not be that of the key established in the first part of the item. And in the left out items the general pattern leads up to a final tonic chord, which is then not played. This system of expectations is one of the baffling parts of music: rhythm, rhyme, and grammar suggest situations that are in some way analogous, but the complete explanation is probably still to be given. At all events we can conclude that the expectation in this case is for a pitch, or a chord-full of pitches, and that these pitches differ by a semitone at least from any that are not expected. Even if the actual pitch forthcoming should differ by a few cycles from that defined by the particular scale it is improbable that any except the most acute will notice it. There is more than one scale known even in western music, and the equally tempered, which has carried the field because of its adaptability to keyboard instruments has no intervals tuned to the simple Pythagorean ratios except the octave. At all events some people will know better what to expect than others. The reader is referred to Franklin's (1956) monograph for a full discussion of tonality.

3. Interval Discrimination. This was a shortened version of Lundin's test. Forty of the fifty items were used, and for this study the items were cast as three- rather than two-choice; as the test is not commercially available on records it was re-recorded on a Hammond organ. The test items ask the subjects to compare the sizes of two melodic intervals. As such it would appear to have little connection with small differences in pitch. Intervals are a most important part of our music, and the ability to judge their size independently of

the notes forming them seems a reasonable ability to expect in a musical person.

2. This was Wing's test 2. The items of this test consist of two chords played on a piano. One of the notes of the second chord may be changed, and the subjects have to say whether the two chords are the same, or whether the note which is changed moves up or down in the second playing of the chord. The difficulty of the chords used ranges from two to four-note chords. The tests are available on tape or record, and for this study the test items were copied directly from a tape.

13. There is clearly a host of techniques available for measuring memory independent of music and pitch. Kelley (1954) devised a battery of them. Of the various types of memory factor that Kelley discusses an auditory test was obviously called for, and a span test seemed more related to music than Rote or Meaningful memory. It was decided to put into the battery an auditory digit span test, as similar to Kelley's test No. 5 as could be judged from his report. In effect the test is an adaptation of the time-honored digit span items of the Binet scale, adapted for group use.

14. Karlin's first study had suggested that there might be little relation between an intellectual task and a test involving sounds, and for that reason there was some doubt that the digit span test would be in the space of the auditory tests to any great extent. Accordingly, another test was sought which would have no obvious relation to pitch or music, but which would involve sounds. Two which seemed possibilities are Seashore's Rhythm and Timbre tests,

with time delays introduced between the first and second stimuli. The timbre test is based on the harmonics of a single fundamental, and as such it involves judging intensities rather than pitches. It was decided to use something like this rather than the rhythm test for two reasons: it was to have had something in common with Test 7--i.e., call for attention through a competing background: and it was technically easier to make. Rather than transcribe the Seashore Timbre or Rhythm test on to tape and then attempt to cut the tape between the first and second stimuli of each item, so as to insert a length of blank tape as a time delay, it was decided to use the harmonics built in on the drawbars of the Hammond organ to supply the different timbres. That is, two slightly different timbres were set upon the drawbars of the instrument's two manuals, intensities were adjusted until judged equal by two or three judges, and records of the two notes were then made on tape. Clearly this is a method which does not supply us with the figures for the energy spectrum that are available for the Seashore test, but it was felt adequate to order a group of subjects. The time delay was from five to seven seconds.

12. This was the same as test 14, but without the time delay.

13. This was Seashore's Tonal Memory test. This test uses the keyboard intervals, and can be expected to that extent to enter into a music factor as well. On the other hand it is atonal, which is to say that a fixed system of pitches is seldom built up; and therefore the musical importance is probably decreased.

The last five tests to be described were designed so that they would require a subject to possess two of the hypothesized abilities for successful performance -- in terms of a factor matrix, load substantially on two of the factors. Western music, as we have seen, is based on a discrete number of tones, and although a small departure from the mathematical frequencies is permitted and often found -- as in, for example, the vibrato -- to diverge too much is an error. For a singer to diverge systematically up or down -- that is, go sharp or flat -- is an important technical fault, no matter how common. The ability to detect slight amounts of sharping or flatting may then be taken as evidence of two things. First, of musicianship, as exhibited by a feeling for, or expectation of, the correct frequencies. Second, of a good ear for pitch, as differences of much less than a semitone will often be found. To do well on these tests, subjects may require either or both abilities.

4. Pitch and music. The items of this test consisted of two notes: the second could be twice the frequency of the first, or somewhat more, or somewhat less than this, and the subjects were asked to say whether the interval was exactly an octave or sharp or flat. An easy item on the test might be 400 c/s followed by 760 c/s, and a difficult item might be 400 c/s followed by 810 c/s. The test was based on the octave, as being the simplest musical interval, and after a pretest based on the fifth had failed.

Ward (1954) examined the relation between the subjective octave and the mathematical, much as in the mel scale of pitch of Stevens and Volkmann (1940), and concluded that the subjective octave was slightly wide of the mathematical.

Lichte (1955) also had found that fifths were seldom tuned to the mathematical ratios, in observations using adjustable oscillators. These effects were, however, expected to be surpassed in size by the variation in ability of the subjects, though it was conceivable that a 400-790 c/s jump might prove to be more often judged short of an octave, than a 400-810 c/s jump be judged wide of an octave.

5. Pitch and music. Here an eight note scale was played, but sometimes did not stay true, but went progressively sharp or flat. If we write down the frequencies of a correct scale as:

360 405 450 480 540 600 675 720 c/s,

then a sharping scale might, for example, be:

360 408 456 489 552 615 693 740 c/s.

6. Pitch and music. This was similar to test 5, except that instead of a scale the item consisted of the first four bars of the tune 'Frere Jacques.' The original reason for putting in both a scale and a tune was again in terms of the one having more within it than simply pitch variation to compete for the hearers' attention.

11. Music and memory. In test 11 sequences of 4, 5, or 6 chords were played, and the subjects had to say whether the first and last chords of a sequence were the same or different. As such it had a similarity to Wing's test of Chord Changes, except that memory is called for, with the first and last chords being separated. Even more, though, it would seem that musicianship would be an asset: the chords were for the most part smooth progressions, and in several of the items the last chord was deliberately approached as a cadence to a tonic, which would give a sense of finality, but would not for

that reason be necessarily the same as the first chord. In standard harmonic notation practice (Piston, 1948) such as item might be V - IV - V - I.

10. Pitch and memory. This test was much like Test 1, except that an interval of about six seconds was put in between the first and second stimuli.

Summary: The prediction was then that there would be a factor structure capable of rotation to something like:

	Factor:	A	B	C
	Test	Music	Pitch	Memory
1.	Pitch, from Seashore	-	x	-
2.	Chord Change, from Wing	x	-	-
3.	Interval Discrimination from Lundin.	x	-	-
4.	Octaves	x	x	-
5.	Scales	x	x	-
6.	Flat or sharp tune	x	x	-
7.	Pitch, masked	-	x	-
8.	Tonality, from Franklin	x	-	-
9.	Bands of noise	-	x	-
10.	Pitch, time delay	-	x	-
11.	Chord sequences	x	-	x
12.	Timbre	-	-	x
13.	Tonal memory, from Seashore	-	-	x
14.	Timbre, time delay	-	-	x
15.	Auditory Digit Span	-	-	x

Tests 4, 5, 6, 7, 10 were made by calculating all required frequencies, recording lengths of tape at each of these at a tape speed of 7 1/2"/sec, and then cutting and splicing required frequencies together. An item from, for example, Test 6 might then have 22 splices in it. Nevertheless this was easier than trying to arrange sharping or flatting by speeding up or slowing down capstan heads. For these tests the tone was in the first instance produced by a Hewlett-Packard oscillator, with a Wave Counter to check the oscillator's accuracy. In practice, when the equipment had warmed up it was frequently possible to record for a minute or more without the oscillator's drifting by a single cycle per second.

Test 15 involved only reading, and an excellent microphone was available for this and all recorded instructions.

Tests 8 and 11 were recorded from a piano.

A master tape of all the tests was then prepared: this included such instructions, sample items, item numbers read aloud at certain points sufficient, it was hoped to avoid subjects' losing the place. From the master tape the working tapes were then copied, and these were then free of splices. For this purpose two machines made by the Ampex Co., were available.

An answer booklet was prepared. One page of introductory remarks was read silently by the subjects: thereafter, for each of the fifteen tests the instructions were on the tape as well as in the booklet. It was felt that as sample items were always given, there should be no occasion for subjects' not comprehending the tasks after each introduction had been read, and the procedure had the advantage of pacing all the subjects alike. The total playing of the tests took somewhat under two hours.

The Experiment.

Subjects. One of the objects of the experiment being the prediction of differences, the services of two rather different subject groups were secured, (i) The Westminster Choir College is a small institution specializing heavily in music, and is the home of the Westminster Choir. Its freshman class in the 1958-59 year was something over seventy students, and this class formed one group of subjects. With one or two incomplete scripts, and absences due to illness, the number of complete records was in the end 67. As the testing was done in early December it can be seen that these students had had about a semester of the Westminster curriculum. This might not sound much time in which to become musicians, but it should be borne in mind that there is an undoubted process of selection to the school. This is partly the school's own screening procedures, but even without this, the curriculum and policy of the institution mean that a large section of the college-entrant population never applies. In fact for the most part only students interested and able in music will want to enter the College. Within this group there will be variation, as always, but some indication of this variation of the students among themselves will be available from the college's own grades. The testing was done in two sessions at the College, an Ampex tape recorder and Speaker-Amplifier providing the source of sound.

(ii) The other group was drawn from an introductory psychology course at Princeton University. There were 35 volunteers, their incentive being extra laboratory credit in their course. In this group a considerable variation in musical aptitude is to be expected. Membership in Princeton does not preclude high musical ability, though it probably means that music is a side interest rather than a prime object of study. In the announcement the freshmen were told that subjects were required for an experiment in hearing, that the testing time would amount to two hours, and the usual arrangements with regard to laboratory credit would apply.

CHAPTER V

Results

The tests were scored, and some three or four of them appeared to have rather skewed distributions, owing to the items being too easy. For this reason the medians were computed, and the cutting scores closest to these were found for all the tests. From the four-fold tables so obtained the tetrachoric correlations were found, rather than product-moment correlations. The matrix of tetrachoric correlations, with the highest side correlation in the diagonal was factored by the method of Principal Axes (Thurstone, 1947), and from an examination of the characteristic roots a judgment was made that three factors accounted for the systematic aspects of the correlation matrix. The factor loadings were found, and an attempt was made to rotate the matrix to simple structure. At this stage it became evident that Test 8, Franklin's test of Tonality had taken up a position in the space that could be explained only with difficulty. It was not discriminating the two groups by much more than 1 1/2 points, with the number of items 25, and appeared to have a negative loading on a factor otherwise identifiable as pitch. The fact that tetrachoric correlations had been used could be invoked to account for the strange position -- with its substantially larger sampling error it was possible that several of the correlations with the other tests were substantially in error, but on further consideration of the nature of the test the idea was conceived that the category Left Out could have been viewed on a number of occasions as wrong: to leave a melody incomplete is certainly not right -- hence might be, on a quick decision, wrong. It was conceivable also that the musical subjects, with their greater knowledge of, and tolerance for, ways in which a tune can work towards its close, would find fewer uses for the incomplete category. It was even possible that with their greater familiarity with change of key, that some of the endings manifestly wrong to a less sophisticated hearer might sound instead as though the intention had been to move into another key. At all events it was decided to rescore the test combining the Wrong, and Left Out categories, and if the device succeeded in giving a more convincing factor matrix without doing violence to the difference in means, it would be its own justification. On this system if a subject had the response Left Out circled, when the test description called for Wrong, it was counted as correct, and vice versa. Circling Right for either of these categories was counted an error, and circling Wrong or Left Out for items coded as Right was also an error. In fact the difference in means does increase very slightly (1.62 as against 1.51) in what is now a test of two rather than three choice items. The variance is low, but it was decided to repeat the factoring with this change. The diagonal estimates for this second factoring were replaced with the communalities available from the first factoring. Table 1 in the Appendix gives the means for the two groups on all the tests, and the means and standard deviations for the total sample.

Table 2 gives the table of tetrachoric intercorrelations for the fifteen tests in the battery.

A listing of the characteristic roots of the table of intercorrelations (with estimated communalities in the diagonal) is given in Table 3-I. From an inspection of these roots three was judged to be the rank of the matrix. Principal Axes factor loadings were computed from the characteristic roots and corresponding characteristic vectors; and appear in Table 3-II.

The matrix of residual correlations was computed, and is given in Table 3-III.

The factor matrix was rotated by the method of Extended Vectors (Thurstone, 1947), and gave the rotated matrix of Table 4-I. The transformation matrix Λ and its various derivatives are quoted in Table 4-II and the sections following. The extended vectors plot is shown in Figure 1.

Interpretation of Factors.

Each factor will be discussed in turn. For the most part loadings below 0.30 will not be considered.

Factor A	Test	4.	Octaves	.34
		5.	Scales	.41
		6.	Flat or sharp tune	.52
		8.	Tonality, from Franklin	.56
		11.	Chord sequences	.40

In each of these tests the stimulus material is based on music, either in being produced by the piano, or being based on a musical structure such as a scale. For these reasons factor A is identified with MUSIC. One other loading should be noted:

Test 3. Interval Discrimination, from Lundin. .00

This loading does not agree with that predicted for it, and the test has appeared on the plot as, in fact, a pitch test. There is no obvious reason why this should be so. The easiest explanation is that with the greater instability of the tetrachoric coefficient the zero correlation with Test 8, Franklin's Tonality, is greatly in error, and that because of this low correlation Tests 3 and 8 must appear orthogonal to each other on the plot. Test 2, Chord Changes, also has a higher loading than the 0.24 on this factor.

Factor B	Test	1.	Pitch, from Seashore.	.66
		2.	Chord Changes, from Wing	.70
		3.	Interval Discrimination from Lundin	.88
		4.	Octaves	.51
		5.	Scales	.33
		6.	Flat or sharp tune	.51
		9.	Bands of Noise	.40
		10.	Pitch, time delay	.35
		13.	Tonal Memory, from Seashore	.46

In each of these tests there are changes in pitch and in six of them the changes are of magnitudes well less than the smallest of the keyboard intervals. For these reasons Factor B has been identified with PITCH. In Tests 2, 3, and 13, the items by their nature could not involve differences smaller than the keyboard intervals: that is, differences many times the differential limen for pitch. It appears that where there is almost any kind of additional musical background the pitch limen is greatly increased--

those people who are poor at pitch discrimination in a pure situation like the Seashore Test will be correspondingly poor in a musical situation. The strong musical components in the factor as it stands suggest that further work might separate out two factors, one on pitch pure and simple, the other a second kind of music factor, though how this would differ from Factor A is uncertain. Test 7, Pitch masked by white noise might have cast some light upon the exact nature of Factor B, but this test was bad in that it was too easy, and its correlations with the rest of the battery are low.

Factor C	Test	5.	Scales	.32
		10.	Pitch, time delay	.60
		12.	Timbre	.38
		13.	Tonal Memory, from Seashore	.65
		14.	Timbre, time delay	.79

Of these tests the three with the high loadings were put in specifically to test for memory effects, and hence Factor C is identified with MEMORY. Test 12, Timbre, was rather easy, and its similarity of form to Test 14, Timbre with time delay, should be borne in mind, as a reason for its having its only large loading on a memory factor. Test 15, Auditory Digit Span had a very low communality with the rest of the battery, and for that reason it could not appear on the extended vectors plot. The suggestion of Karlin's (1941) work is to this extent confirmed, that an intellectual task has little to do with the auditory domain. It is of interest that this test gives a substantial mean difference in favor of the Princeton group.

The factor structure is a partial confirmation of the (1941) work of Karlin. We have noted that his Tonal Sensitivity was defined by a pitch test and an interval discrimination test. This study has added other pitch tests as a check. In this study there is also an interval discrimination test which loads on the factor, plus two other tests using musical material. These are Wing's Chord Changes and the Seashore Tonal Memory.

The Memory factor has also been confirmed.

The single Music factor in this study is considerably less correlated with Pitch than the Retentivity factor was correlated with Tonal Sensitivity in Karlin's (1941) study, and has its largest loadings on tests which were not available at the time of Karlin's study. It is therefore plausible that Karlin's concept of Retentivity can now be subdivided into music and pitch. As such there need now be no great disagreement with conclusions such as Wing's (1941). It will be recalled that he described a large general factor with smaller bipolars. The large general factor would, in fact, be our music factor--which would also agree with the finding of Franklin's (1956) first factor matrix.

The existence of three factors, of pitch, music, and memory being indicated, let us now examine the effect that they will have on the validation problem. The two groups of subjects having been selected such that differences in musical aptitude were highly probable, we may then assess the efficiency of the testing by attempting to separate out the two groups on the basis of their scores. They are two ways at least of doing this.

By the Linear Discriminant Function, (Hoel, 1947) a weighted sum of the scores is found for each subject such that in the resultant listing the two groups of subjects are maximally discriminated. Hoel gives a geometric analogy.

Secondly we can produce an arbitrary criterion by giving each of the first group a score of one, and each of the second group a score of zero, and then attempt to predict this criterion by the methods of multiple regression. It can be shown that the two methods produce the same weights for the predictors, or at all events a simple multiple of them, but in addition the multiple regression will enable us to find the multiple correlation, from which we may judge the efficiency of the discrimination achieved.

Any number of predictors may be used. Rather than use all fifteen test scores we shall use the results of the factor analysis to reduce these to three. This will save computational bulk, and makes sense once we have judged the rank of the correlation matrix to be three.

Let us define certain matrices. The order of each is quoted after the definition, along with a statement of what the rows and columns stand for.

S Scores of subjects on tests, in standard scores.
i for tests, by j for subjects.

F Factor matrix of tests, from Principal Axes factoring.
i for tests, by r for factors.

P Factor scores of subjects.
i for tests, by j for subjects.

Λ Rotation of the F matrix to simple structure.

r by r, for factors.

V Rotated Factor matrix, i.e. $F\Lambda = V$

i for tests by r for factors.

X Factor scores obtained from rotated matrix V

i for tests by j for subjects.

Then: by the fundamental theorem of factor analysis

$$S = FP \quad .$$

Rearranging,

$$P = (F'F)^{-1} F'S \; ,$$

from which P may be calculated.

Also let X be defined by

$$S = VX \quad .$$

Then

$$X = (V'V)^{-1} V'S \; .$$

The relation of X to P may be found as follows:

$$S = FP = VX \quad .$$

Substituting for V

$$FP = F\Lambda X \quad ,$$

from which

$$X = \Lambda^{-1}P \quad .$$

It is more useful with machine computation to compute X directly from S by substituting for P:

$$X = \Lambda^{-1}(F'F)^{-1} F'S \; .$$

P and X are then representations, in three rather than fifteen scores, of the battery.

The multiple regression calculation is summarized in Table 5, which gives the multiple correlation, and the weights for the three factor scores. The calculation was performed using both the P and the X matrices of factor scores as predictors, since it was possible to use the same program on the computer. The multiple correlation is the same in the two cases, because the transformation Λ has neither gained nor lost information, which was also a check on the computations. The weights in the case of the rotated factor matrix are easier to interpret. The large weight is the first, and has reference to the column of the factor matrix which was identified with music. These weights mean that those tests which load highest on the music Factor A are the most effective predictors--in general numbers 2, 4, 5, 6, 8, and 11. These are the tests which have the largest mean differences between the two groups. The other two factors, which have been identified with pitch and memory have positive weights, but smaller. The tests which load on Factors B and C are therefore of less effect in discriminating the two groups.

If we proceed to compute the predicted score we arrive at the listing of Table 6-I. In that table an asterisk after any given score indicates a member of the Princeton group. There is a notable concentration of Princeton scores at the low end, plus a smaller number considerably higher, and the implication is that these would have most ability in music. In Table 6-I an arbitrary cutting point has been indicated by the solid line at the 73rd score: a frequency count of the two groups about this point gives the fourfold Table 6-II, as a short summary of the discrimination achieved.

We have the foregoing on the assumption that the Choir College students had more ability in the area than the Princeton. The results set forth suggest that this is reasonable. Obviously the students in any one institution will differ among themselves in ability. The multiple regression analysis that we have run does not take this into account. For the Westminster Choir College group an external criterion was available in the college's own grades, to which the author has access. There is a risk with such a criterion, that no useful confirmation is going to emerge if this independent measure should happen to have nothing in common with the predictors. The ideal criterion, which in this case might be an assessment of the actual effectiveness in various situations of the subjects, is seldom available, least of all while they are college freshmen.

In fact the measure consisted of a 120-item test covering aspects of eartraining traditional in music--telling sizes of intervals, types of chords and so on. As such it would not involve any use of pitches smaller than the keyboard intervals. Also it involved domains which this investigation has not attempted to cover, such as rhythm. The multiple regression calculations are summarized in Table 7. The multiple correlation of 0.53 is substantially less that the 0.68 of the previous section--an effect that has often enough been encountered in selected populations. Also the largest weight is now no longer on the music factor, but on the pitch, though in fact all three weights are now closer to each other than in the previous analysis. Many tests load on this pitch factor, including three that were thought to involve music more than pitch. To have named Factor B pitch is for that reason probably an oversimplification.

CHAPTER VI

Summary and Conclusions

We attempt now to answer some of the questions posed at the beginning of the experiment.

1. The factor structure suggests factors corresponding to pitch, music and memory which are slightly correlated. Of these, the pitch factor has been noted before by Karlin (1942), Franklin (1956) and others. However we have shown that several musical tests also load upon this pitch factor. This was suggested in Karlin's (1941) work as a Tonal Sensitivity factor defined by two tests only. Some of the musical tests involved intervals less than semitones, and their intimate connection with a pitch factor is plausible, but others are based only on the keyboard intervals. This result suggests that those who cannot hear very small differences in pitch in a rather pure situation can often not hear very much larger differences in pitch also if there is any sort of additional structure in the stimuli. This possibility, which was not part of the original plan of the study, calls for additional investigation.
2. A Memory factor has also been found by previous investigators. (Karlin, 1942; Franklin, 1956). Pitch and memory being different effects, it is evident that account should be taken of the possible effects of memory in measures of pitch discrimination.
3. We confirm the finding of Karlin (1941) and others that memory in the digit-span intelligence-type test has little or nothing in common with the auditory domain.

4. A Music factor has been found by some investigators (Wing, 1941; Franklin, 1956), and suggested by others (Karlin, 1941; Drake, 1939). This study, with a group of music tests collected from several sources, has given confirmation of this factor.
5. None of the factor studies considered has attempted to assess the value in a predictive sense of the factors found. The treatment of the results here was to summarize the fifteen test scores for each subject into three factor scores. The factor scores were computed from the rotated factor matrix, and so they preserve the character of the three factors defined by that rotation. A subject can then be spoken of as having a high or low score on music, memory or pitch, meaning, in effect, high scores on that group of tests underlying the factor. Subjects were tested from two institutions. In the one they were all following a musical curriculum; in the other music can be the object of study, but is more probably only an interest. Using the factor scores as predictors of the institution of the various subjects, a multiple correlation of 0.68 was found. The largest weight of these three is that of the Music factor score. This therefore confirms the importance of the music-type tests as against the more physical type test. Within the musical group of students, and against a criterion of an ear-training test in music, the multiple correlation was 0.53, and the most important predictor now becomes the pitch factor score. It is precisely those tests composed especially for this study, which were designed to bridge the borderline between pitch and music, which load on both the pitch and the music factors,

which give large mean differences between the two groups of subjects, and which are of most importance therefore in predicting both the institution of any one subject, and also the standing of the subjects in the musical group. These tests are therefore viewed as one of the main contributions to testing of musical aptitude of this study.

REFERENCES

Aliferis, J., & Stecklein, J. E. Measurement of music achievement at college entrance. J. Appl. Psychol., 1955, 39, 263-272.

Bachem, A. Time factors in relative and absolute pitch determination. J. Acoust. Soc. Amer., 1954, 26, 751-753.

Brown, A. W. The reliability and validity of the Seashore Tests of Musical Talent. J. Appl. Psychol., 1928, 12, 468-476.

Drake, R. M. Four new tests of musical talent. J. Appl. Psychol., 1933, 17, 136-147.

Drake, R. M. Factorial analysis of music tests by the Spearman Tetrad-Difference Technique. J. Musicol., 1939, 1, 6-16.

Farnsworth, P. R. Are music capacity tests more important than intelligence tests in the prediction of several types of music grades? J. Appl. Psychol., 1935, 19, 347-350.

Farnsworth, P. R. The social psychology of music. New York: Dryden, 1958.

Franklin, E. Tonality as a basis for the study of musical talent. Göteborg: Gumperts Forlag, 1956.

Harris, J. D. A search toward the primary auditory abilities. Memorandum Report No. 57-4. BuMed Project NM 22 01 20.2.1. Medical Research Lab., U. S. Naval Submarine Base, New London, Connecticut, 1957.

Hoel, P. G. Introduction to mathematical statistics. New York: Wiley, 1947.

Karlin, J. E. Music ability. Psychometrika, 1941, 6, 61-65.

Karlin, J. E. A factorial study of auditory function. Psychometrika, 1942, 7, 251-279.

Karlin, J. E.,& Stevens, S. S. Auditory tests for the ability to discriminate the pitch and loudness of noises. Amer. Psychologist, 1946, 1, 238.

Kelley, H. P. A factor analysis of memory ability. Doctoral Thesis, Princeton University, 1954.

König, E. Effect of time on pitch discrimination thresholds under several psychophysical procedures: Comparison with intensity discrimination thresholds. J. Acoust. Soc. Amer., 1957, 29, 606-612.

König, H. Über das musikalische Gedächtnis. Zeitschr. f. Psychologie, 1928, 108, 398-420.

Kwalwasser, J. Exploring the musical mind. New York: Coleman-Ross, 1955.

Lichte, W. H. One man's preferred fifth... Amer. J. Psychol., 1955, 68, 312-315.

Lowery, H. Cadence and phrase tests in music. Brit. J. Psychol., 1926, 17, 111-118.

Lowery, H. Music Memory. Brit. J. Psychol., 1929, 19, 397-404.

Lundin, R. W. The development and validation of a set of musical ability tests. Psychol. Monogr., 1949, 63:305, 1-20.

Lundin, R. W. An objective psychology of music. New York: Ronald Press, 1953.

McLeish, J. The validation of Seashore's measures of musical talent by factorial methods. Brit. J. Psychol., Stat. Sect., 1950, 3, 129-140.

Neu, D. M. A critical review of the literature on "Absolute Pitch." Psychol. Bull., 1947, 44, 249-266.

Piston, W. Harmony. (Rev. ed.) New York: W. W. Norton, 1948.

Revesz, G. Introduction to the psychology of music. Norman: Univer. of Oklahoma Press, 1954.

Schoen, M. Tests of musical feeling and understanding. J. Comp. Psychol., 1925, 5, 31-52.

Seashore, C. E. The psychology of music. New York: McGraw-Hill, 1938.

Seashore, C. E., Lewis, D.,& Saetveit, J. G. Seashore Measures of musical talents. (Revised.) New York: Psychological Corporation, 522 Fifth Avenue, New York 36, N.Y., 1956.

Stevens, S. S., & Volkmann, J. The relation of pitch to frequency: a revised scale. Amer. J. Psychol., 1940, 53, 329-353.

Thurstone, L. L. Multiple factor analysis. Chicago: Univer. of Chicago Press, 1947.

Wallach, H., Newman, E. B., & Rosenzweig, M. R. The precedence effect in sound localization. Amer. J. Psychol., 1949, 62, 315-336.

Ward, W. D. Subjective musical pitch. J. Acoust. Soc. Amer., 1954, 26, 369-380.

Wing, H. D. A factorial study of musical tests. Brit. J. Psychol., 1941, 31, 341-355.

Wing, H. D. Tests of musical ability and appreciation. Brit. J. Psychol. Monogr. Suppl., 1948, No. 27.

Appendix I

Tables 1-7

Figure 1

Table 1

	Test	Mean Scores			Standard Deviation
		Westminster	Princeton	Overall	Overall
1.	Pitch, Seashore	30.25	29.28	29.92	3.488
2.	Chord Change, Wing	26.75	21.91	25.09	4.499
3.	Interval Discrimination, Lundin.	25.85	22.23	24.61	5.400
4.	Octaves	20.77	15.83	19.08	4.283
5.	Scales	16.36	12.54	15.05	3.882
6.	Flat or sharp tune	19.55	14.97	17.98	4.073
7.	Pitch, masked	34.43	32.31	33.67	2.784
8.	Tonality, Franklin	21.25	19.63	20.70	1.959
9.	Bands of noise	30.04	30.42	30.18	2.398
10.	Pitch, time delay	33.82	31.46	33.01	3.117
11.	Chord sequences	21.55	19.20	20.75	3.035
12.	Timbre	33.46	32.97	33.29	1.978
13.	Tonal Memory, Seashore	28.63	25.29	27.48	3.189
14.	Timbre, time delay	31.86	30.03	31.23	3.040
15.	Auditory digit span	6.52	7.91	7.00	2.072

Table 2

Tetrachoric Intercorrelations of Test Battery.

Test:	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1		.53	.58	.25	.43	.46	.32	-.03	.43	.40	.15	.11	.44	.13	-.04
2			.76	.69	.42	.72	.42	.15	.26	.50	.49	.27	.67	.42	.09
3				.62	.38	.53	.22	-.02	.33	.33	.28	.02	.46	.05	.21
4					.62	.62	.18	.13	.25	.54	.43	.22	.73	.32	-.09
5						.53	.38	.28	.22	.55	.44	.20	.62	.38	.00
6							.36	.32	.10	.51	.40	.00	.46	.17	-.03
7								.15	.15	.42	.35	.11	.32	.32	-.06
8									.12	.03	.20	.18	.06	.03	-.33
9										.25	.21	.17	.31	-.05	-.02
10											.22	.18	.73	.57	-.05
11												.11	.38	.22	.07
12													.34	.39	.02
13														.60	.04
14															.07
15															

Table 3-I

Characteristic Roots of Matrix of Intercorrelations

(Estimated Communalities inserted in diagonal cells.)

1.	5.37670
2.	1.15119
3.	.75949
4.	.48700
5.	.42127
6.	.35044
7.	.15538
8.	.05803
9.	.05562
10.	.02188
11.	-.05372
12.	-.10021
13.	-.18965
14.	-.26466
15.	-.27157

Table 3-II

Principal Axes Factor Loadings.

Test	Factor I	Factor II	Factor III
1.	.583	-.307	.184
2.	.860	-.171	.045
3.	.679	-.538	.224
4.	.778	-.029	-.085
5.	.708	.112	-.182
6.	.721	-.215	-.302
7.	.465	.090	-.112
8.	.188	.064	-.526
9.	.355	-.209	.062
10.	.730	.284	.129
11.	.509	-.009	-.242
12.	.278	.304	.048
13.	.842	.252	.188
14.	.503	.617	.191
15.	.015	-.074	.316

Table 3-III

Matrix of Residual Intercorrelations after Extraction of Three Principal Axes Factors

Test:	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1		-.03	.02	-.09	.00	.03	.10	-.03	.15	.04	-.10	.03	-.01	-.01	-.13
2			.07	.02	.16	.08	.04	-.02	-.08	-.08	.06	.08	-.02	.08	.05
3				.10	.00	-.01	-.02	.00	-.04	-.04	-.02	-.02	-.02	.00	-.03
4					.06	.03	-.19	-.06	-.03	-.01	.01	.02	.10	-.04	-.07
5						-.01	.02	.04	.00	.02	.04	-.02	.03	-.01	.01
6							.01	.04	-.18	.08	-.04	-.12	-.04	.00	.04
7								.00	.01	.07	.09	-.04	-.07	.05	-.02
8									.10	-.06	-.02	.13	-.02	.00	-.16
9										.04	.04	.13	.05	-.11	-.06
10											-.12	-.11	.02	.00	.02
11												-.02	.00	.02	.14
12													.02	.05	.02
13														.01	-.01
14															.05
15															

Table 4-I

Rotated Factor Matrix

Test	A Music	B Pitch	C Memory
1. Pitch, Seashore	.012	.655	.102
2. Chord Change, Wing	.237	.700	.268
3. Interval Discrimination, Lundin	.000	.884	-.017
4. Octaves	.336	.506	.284
5. Scales	.408	.332	.319
6. Flat or sharp tune	.518	.511	.017
7. Pitch, masked	.261	.210	.226
8. Tonality, Franklin	.560	-.106	-.102
9. Bands of noise	.054	.395	.023
10. Pitch, time delay	.126	.346	.602
11. Chord sequences	.396	.257	.110
12. Timbre	.053	.003	.381
13. Tonal Memory, Seashore	.106	.463	.652
14. Timbre, time delay	.000	.000	.788
15. Auditory digit span	-.295	.172	.090

Table 4-II

Transformation Matrix

	A	B	C
I.	.330	.665	.443
II.	.023	-.654	.777
III.	-.944	.360	.447

Table 4-III

$\Lambda'\Lambda$

Correlations of Reference Vectors

	A	B	C
A	1.0000	-.1354	-.2579
B		1.0000	-.0526
C			1.0000

Table 4-IV

$(\Lambda'\Lambda)^{-1}$

	A	B	C
A	1.0974	.1640	.2917
B		1.0273	.0964
C			1.0803

Table 4-V

$D(\Lambda'\Lambda)^{-1}D$

Correlations of Primary Factors

	A	B	C
A	1.0000	.1545	.2679
B		1.0000	.0915
C			1.0000

Table 4-VI

Λ^{-1}

	I	II	III
A	.6001	.1445	-.8460
B	.7803	-.5935	.2583
C	.6390	.7832	.2425

Table 5

Regression weights for predicting College attended

Principal Axes Factors	Weight
I	.4892
II	.0631
III	-.4121

Rotated Axes Factors	Weight
A: Music	.6136
B: Pitch	.1312
C: Memory	.1000

Multiple Correlation = 0.68

Table 6-I

Predicted scores for best discrimination of two groups.

A score obtained by a Princeton student is indicated with an asterisk.

1.	5.9531
2.	5.8267
3.	5.3358
4.	5.3154
5.	5.2319
6.	4.7631
7.	4.2498
8.	4.1368
9.	4.1210
10.	3.9513
11.	3.8416
12.	3.7984
13.	3.7630
14.	3.7303
15.	3.6315
16.	3.4954
17.	3.4581
18.	3.2713
19.	3.2471
20.	3.2210
21.	3.2174
22.	3.2012
23.	2.8552
24.	2.6780
25.	2.6691
26.	2.6030
27.	2.4835
28.	2.4132
29.	2.2966*
30.	2.2051
31.	2.1506
32.	1.9146
33.	1.5800
34.	1.5554
35.	1.3783
36.	1.3146
37.	1.2268
38.	1.0476
39.	1.0412
40.	0.9167
41.	0.9097
42.	0.8785
43.	0.8297
44.	0.7549*
45.	0.6997
46.	0.6191
47.	0.5243*
48.	0.3027
49.	0.2015*
50.	0.1902*
51.	0.0897
52.	0.0523
53.	0.0218
54.	-0.0242*
55.	-0.0272*
56.	-0.0747
57.	-0.1552
58.	-0.2214
59.	-0.2276
60.	-0.2795
61.	-0.3031
62.	-0.5191*
63.	-0.5347
64.	-0.5887
65.	-0.8339
66.	-1.0781*
67.	-1.1497
68.	-1.1521
69.	-1.2763
70.	-1.3787*
71.	-1.3811
72.	-1.4175
73.	-1.4766
74.	-1.4940*
75.	-1.6381
76.	-1.7536*
77.	-1.8677
78.	-1.8911*
79.	-1.9009
80.	-2.1907*
81.	-2.3802*
82.	-2.4045*
83.	-2.4521
84.	-2.5800*
85.	-2.9080*
86.	-2.9802*
87.	-3.1804*
88.	-3.4918*
89.	-3.9690*
90.	-4.4341*
91.	-4.7015*
92.	-4.7948*
93.	-4.8289*
94.	-5.0589*
95.	-5.1195*
96.	-5.6029*
97.	-5.7864*
98.	-6.3836*
99.	-6.7242*
100.	-7.1461*
101.	-7.8002*
102.	-9.7056*

Table 6-II

Degree of Discrimination achieved in predicting College attended.

	Westminster Choir College	Princeton University
Above Cutting Score	63	10
Below Cutting Score	4	25

Table 7

Multiple regression of three factor scores of Westminster Choir College students on Westminster ear-training criterion.

Rotated Axes Factors	Regression Weights
A: Music	.2356
B: Pitch	.6855
C: Memory	.3875

Multiple Correlation = 0.53

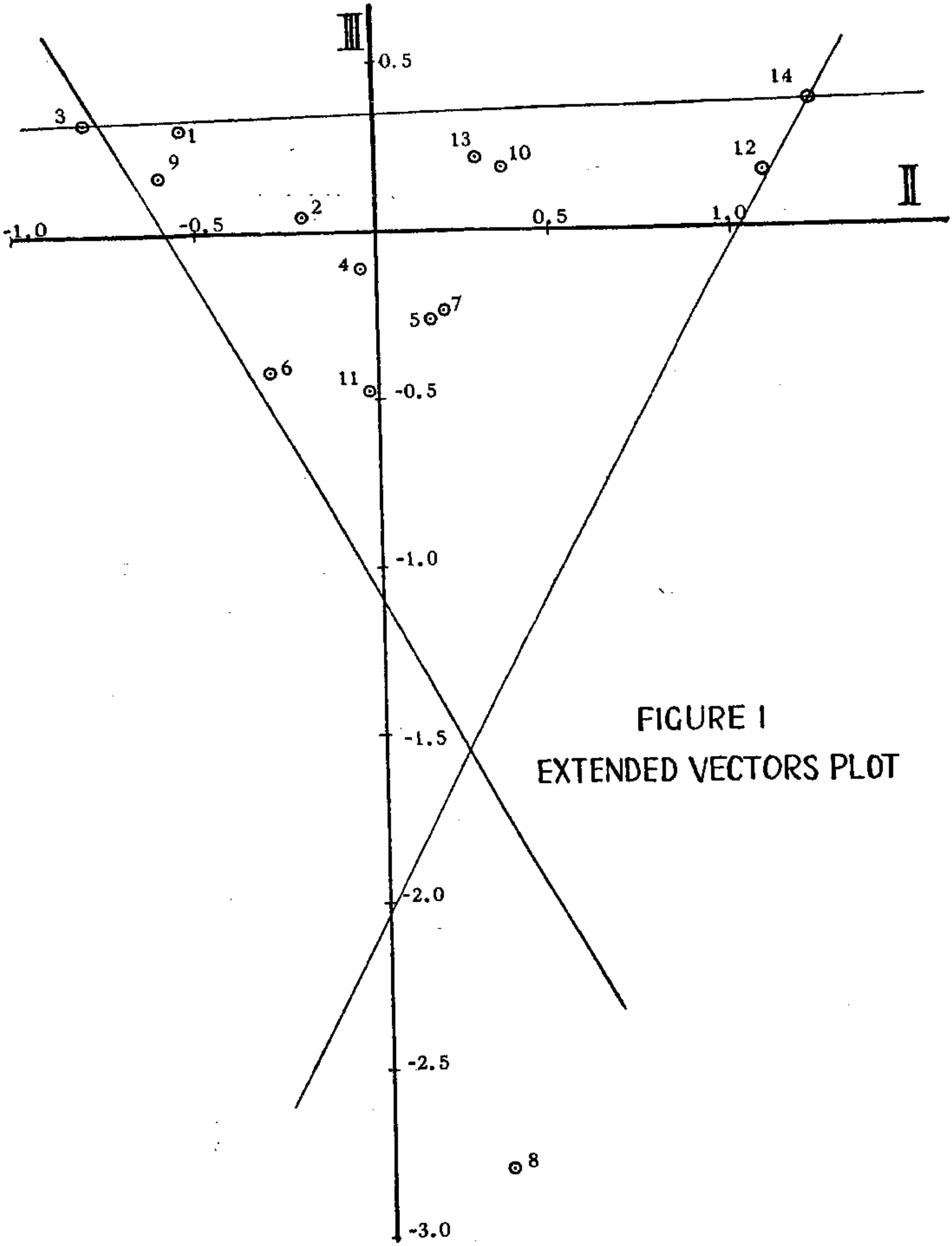

FIGURE I
EXTENDED VECTORS PLOT

Appendix II

The Test Booklet

TESTS IN PITCH PERCEPTION

NAME

INSTITUTION

B. D. Faulds
Educational Testing Service
Princeton, N. J.

November, 1958

GENERAL DIRECTIONS

You are taking part in an experiment designed to investigate certain aspects of hearing. You will hear the test items played over a loudspeaker, and for every item there is a place in this answer booklet where you will record your answer.

Each section is explained for you as you come to it, and there are practice items before each test actually begins. These explanations are both read over the tape, and printed in your answer booklet, so that you can follow them in front of you as they are read out.

If, after the instructions have been read, and you have tried the practice items, you are still not sure what you are supposed to do, then raise your hand, and the experimenter can then play the instructions over again, or else answer your questions himself.

In most of the tests the items are easy to begin with, becoming more difficult. Even though you should feel some of the items too difficult, do not skip them. Give the best answer you can. Answering every item gives you the best chance for a good score on any test.

You do not have much time to answer any one item, so do not delay making your decision, or you may miss the next item.

Now look over a few of the following pages, to see what sort of instructions are given, and if you have any questions, ask them now.

Turn now to the first test, and prepare for the start of the recordings.

TEST 1

Listen to practice item A:

H L

The second note was higher than the first. Circle the "H" in Item A.

Listen to item B:

H L

The second note was lower than the first, and you would circle "L" in Item B.

Each item in the test will consist of two notes, and you are to decide whether the second note is Higher or Lower than the first.

If the second note is Higher, circle "H" on your answer sheet.

If the second note is Lower, circle "L" on your answer sheet.

Here is practice item C for you to try:

Item C: H L

The second note was lower and you should have circled "L".

Now try practice item D:

Item D: H L

The second note was higher, and you should have circled "H".

Now try the test. In this test the number of every fifth item is read out to help you keep the place. Mark every item. If you are not sure of any item use your best judgment; but do not leave blanks.

Turn the page and get ready.

TEST 1

1.	H	L
2.	H	L
3.	H	L
4.	H	L
5.	H	L
6.	H	L
7.	H	L
8.	H	L
9.	H	L
10.	H	L
11.	H	L
12.	H	L
13.	H	L
14.	H	L
15.	H	L
16.	H	L
17.	H	L
18.	H	L
19.	H	L
20.	H	L
21.	H	L
22.	H	L
23.	H	L
24.	H	L
25.	H	L
26.	H	L
27.	H	L
28.	H	L
29.	H	L
30.	H	L
31.	H	L
32.	H	L
33.	H	L
34.	H	L
35.	H	L

TEST 2

Listen to these two chords:

D S U

id you think they were the same? Well, they were; and if you were to ircle "S" for Same, your answer would be correct.

Now try this:

D S U

ne of the notes of the chord went Down, and you would have to circle "D" 'or down.

Here is a more difficult one, as the note change is in the middle f the others:

D S U

The note change went Up that time, and had you circled "U" for up, your answer would be correct.

NOW TRY THE TEST.

Ready:

1.	D	S	U	11.	D	S	U	21.	D	S	U
2.	D	S	U	12.	D	S	U	22.	D	S	U
3.	D	S	U	13.	D	S	U	23.	D	S	U
4.	D	S	U	14.	D	S	U	24.	D	S	U
5.	D	S	U	15.	D	S	U	25.	D	S	U
6.	D	S	U	16.	D	S	U	26.	D	S	U
7.	D	S	U	17.	D	S	U	27.	D	S	U
8.	D	S	U	18.	D	S	U	28.	D	S	U
9.	D	S	U	19.	D	S	U	29.	D	S	U
10.	D	S	U	20.	D	S	U	30.	D	S	U

TEST 3

The name of the next test is Interval Discrimination. In music an interval is the distance between any two notes on the scale. The object of this test is to measure how well you can distinguish various intervals. In each item of the test you will hear a pair of notes played for the first interval, followed by another pair of notes — the second interval.

Here is practice item A. You should circle:

L if you think the second interval is larger than the first,

= if you think the two intervals are equal,

S if you think the second interval is smaller than the first.

Here is the item:

L = S

The second interval was larger, and so we should have circled the "L".

When two intervals are the same the number of steps on the scale which lie between the first and second notes and the third and fourth notes will be the same. This does not mean that the ACTUAL notes are the same, but the intervals are.

Here is practice item B. Again circle:

L if you think the second interval is larger than the first,

= if you think the two intervals equal,

S if you think the second interval smaller than the first.

Here is the item:

L = S

Those were equal intervals. You should have circled the " = ".

Now try practice item C:

L = S

You should have circled "S", because the second interval was smaller.

In the first twenty items the second note of each interval is higher.

Turn the page and get ready for the first item.

TEST 3

1.	L	=	S	11.	L	=	S
2.	L	=	S	12.	L	=	S
3.	L	=	S	13.	L	=	S
4.	L	=	S	14.	L	=	S
5.	L	=	S	15.	L	=	S
6.	L	=	S	16.	L	=	S
7.	L	=	S	17.	L	=	S
8.	L	=	S	18.	L	=	S
9.	L	=	S	19.	L	=	S
10.	L	=	S	20.	L	=	S

In the second twenty items the second note of each interval is lower.

Here are three practice items.

Item A: L = S

You should have circled the " = ", as they were equal intervals.

Item B: L = S

Here you should have circled the "S", as the second interval was smaller.

Item C: L = S

These were again equal intervals.

Now turn the page and try the second twenty items.

TEST 3 (continued)

21.	L	=	S
22.	L	=	S
23.	L	=	S
24.	L	=	S
25.	L	=	S
26.	L	=	S
27.	L	=	S
28.	L	=	S
29.	L	=	S
30.	L	=	S
31.	L	=	S
32.	L	=	S
33.	L	=	S
34.	L	=	S
35.	L	=	S
36.	L	=	S
37.	L	=	S
38.	L	=	S
39.	L	=	S
40.	L	=	S

TEST 4

Listen to these two notes:

They are exactly an octave apart. Here are two similar notes, but listen for a slight difference:

The second note was too high for it to be a true octave any longer.

Or in these two notes:

The second note was too low, and the interval between the notes was less than an octave.

Now in the test items you should:

circle H if the second note is too high for the interval to be an exact octave,
circle C for correct, if the second note is the exact octave of the first,
circle L if the second note is too low for the interval to be an octave.

Here is practice item A:

H C L

The second note was too low, the interval less than an octave, and you should have circled "L".

Here is practice item B:

H C L

The second note was exactly an octave above the first, and you should have circled the "C", for correct.

The number of every fifth item will be read, to help you keep the place.

Now turn the page, and get ready for the test items.

TEST 4

1. H C L
2. H C L
3. H C L
4. H C L
5. H C L
6. H C L
7. H C L
8. H C L
9. H C L
10. H C L
11. H C L
12. H C L
13. H C L
14. H C L
15. H C L
16. H C L
17. H C L
18. H C L
19. H C L
20. H C L
21. H C L
22. H C L
23. H C L
24. H C L
25. H C L

TEST 5

You have just tried to judge whether a single note was sharp or flat.

Now listen to this:

It was an eight note scale, but it went progressively too high, and finished sharp.

We are going to play such scales to you, and they may finish too high, too low, or in tune.

Circle H if a scale finishes too high,

Circle C if a scale finishes correctly,

Circle L if a scale finishes too low.

Here is a practice item:

H C L

It went flat, so "L" for low would be your answer.

Now try the test: the number of every fifth item is read out to help you keep the place.

BE SURE TO ANSWER EVERY ITEM. If you are not sure of the correct answer use your best judgment.

Turn the page and get ready for the test items.

TEST 5

1.	H	C	L
2.	H	C	L
3.	H	C	L
4.	H	C	L
5.	H	C	L
6.	H	C	L
7.	H	C	L
8.	H	C	L
9.	H	C	L
10.	H	C	L
11.	H	C	L
12.	H	C	L
13.	H	C	L
14.	H	C	L
15.	H	C	L
16.	H	C	L
17.	H	C	L
18.	H	C	L
19.	H	C	L
20.	H	C	L
21.	H	C	L
22.	H	C	L
23.	H	C	L
24.	H	C	L
25.	H	C	L

TEST 6

This test is very much like the one you have just finished, only here we have the beginning of a little tune which may go sharp, flat, or end on pitch.

Again you are to circle H if it ends too <u>high</u>,
C if it ends <u>correctly</u>,
L if it ends too <u>low</u>.

Now try the two practice items.

Item A: H C L

Item B: H C L

The first kept in tune, so you should have circled "C", for correct.

The second finished too low, so you should have circled the "L".

In the test the number of every fifth item will be read to help you keep the place. BE SURE TO ANSWER EVERY ITEM. If you are not sure of the correct answer use your best judgment.

Ready:

1.	H	C	L	11.	H	C	L	21.	H	C	L
2.	H	C	L	12.	H	C	L	22.	H	C	L
3.	H	C	L	13.	H	C	L	23.	H	C	L
4.	H	C	L	14.	H	C	L	24.	H	C	L
5.	H	C	L	15.	H	C	L	25.	H	C	L
6.	H	C	L	16.	H	C	L				
7.	H	C	L	17.	H	C	L				
8.	H	C	L	18.	H	C	L				
9.	H	C	L	19.	H	C	L				
10.	H	C	L	20.	H	C	L				

TEST 7

In this test each item will consist of two notes played against a background of noise. You are to decide in each case whether the second note is higher or lower than the first.

You are to circle H if the second note is higher,
L if the second note is lower.

Listen now to practice item A:

H L

The second note was Higher than the first, so you should have circled "H".

Here is practice item B:

H L

The second note was Lower, so you should have circled "L".

As you can see, the task is much like one of the earlier tests. In the test the number of every fifth item will be read to help you keep your place. BE SURE TO ANSWER EVERY ITEM. If you are not sure of the correct answer, use your best judgment.

Ready:

1. H L
2. H L
3. H L
4. H L
5. H L
6. H L
7. H L
8. H L
9. H L
10. H L
11. H L
12. H L
13. H L
14. H L
15. H L
16. H L
17. H L
18. H L
19. H L
20. H L
21. H L
22. H L
23. H L
24. H L
25. H L
26. H L
27. H L
28. H L
29. H L
30. H L
31. H L
32. H L
33. H L
34. H L
35. H L

TEST 8

Listen to this little tune:

It probably sounded incomplete to you, as though the last note had been left out.

Listen to another little tune:

This one probably sounded as though the last note was wrong.

Here is a third little tune:

The last note of this one probably sounded right.

In a melody the last note or chord is of some importance in giving a sense of finality or character to the tune. The items of this test will be short melodies like those we have played, and you should consider their endings.

Circle the word Right if you think the last note was right,

Circle the word Wrong if you think the last note was wrong,

Circle the words Left out if you think the last note was left out.

Here is a practice item:

Right Wrong Left out

There the last note was left out.

Here is another practice item:

Right Wrong Left out

There the last note was wrong.

It should be added that a wrong note need not always sound as a strong discord. It might just as easily be only a slight change from what is expected. ANSWER EVERY ITEM. If you are not certain, give your best judgment.

Now turn the page, and get ready to try the test.

TEST 8

1. Right Wrong Left out
2. Right Wrong Left out
3. Right Wrong Left out
4. Right Wrong Left out
5. Right Wrong Left out
6. Right Wrong Left out
7. Right Wrong Left out
8. Right Wrong Left out
9. Right Wrong Left out
10. Right Wrong Left out
11. Right Wrong Left out
12. Right Wrong Left out
13. Right Wrong Left out
14. Right Wrong Left out
15. Right Wrong Left out
16. Right Wrong Left out
17. Right Wrong Left out
18. Right Wrong Left out
19. Right Wrong Left out
20. Right Wrong Left out
21. Right Wrong Left out
22. Right Wrong Left out
23. Right Wrong Left out
24. Right Wrong Left out
25. Right Wrong Left out

TEST 9

Listen to these two sounds; try to decide whether the second is higher or lower than the first:

They were noises, but they were not identical, and you might well conclude that the second sounded higher than the first.

We do not think of noises as having any particular pitch, since any noise is a mixture of pitches. But if one noise does have higher pitches within it, then it should sound higher.

Here are two practice items. Each will consist of two noises, and you should try to decide which noise has the higher pitch.

Circle H if you think the second noise has the higher pitch,

Circle L if you think the second noise has the lower pitch.

Here is item A:

H L

The second noise was higher, and you should have circled "H".

Here is item B:

H L

The answer to this was "L", since the second noise was lower.

The number of every fifth item will be read to help you keep the place.

Now turn the page and get ready to try the test.

TEST 9

1.	H	L
2.	H	L
3.	H	L
4.	H	L
5.	H	L
6.	H	L
7.	H	L
8.	H	L
9.	H	L
10.	H	L
11.	H	L
12.	H	L
13.	H	L
14.	H	L
15.	H	L
16.	H	L
17.	H	L
18.	H	L
19.	H	L
20.	H	L
21.	H	L
22.	H	L
23.	H	L
24.	H	L
25.	H	L
26.	H	L
27.	H	L
28.	H	L
29.	H	L
30.	H	L
31.	H	L
32.	H	L
33.	H	L
34.	H	L
35.	H	L

TEST 10

In this test we again ask you to decide which of two notes is the higher. Here there is a time delay. One note will be sounded, and then there will be a gap of from five to seven seconds before the next is sounded.

If you think the second note is <u>higher</u> than the first circle "H" for your answer.

If you think the second note is <u>lower</u> than the first circle "L" for your answer.

If you are not sure of any item do not leave it out, but make your best judgment, and fill that in.

Here is a practice item:

Item A: H L

The answer was "H", as the second note was higher.

Now try the test.

Ready:

1.	H	L	13.	H	L	25.	H	L
2.	H	L	14.	H	L	26.	H	L
3.	H	L	15.	H	L	27.	H	L
4.	H	L	16.	H	L	28.	H	L
5.	H	L	17.	H	L	29.	H	L
6.	H	L	18.	H	L	30.	H	L
7.	H	L	19.	H	L	31.	H	L
8.	H	L	20.	H	L	32.	H	L
9.	H	L	21.	H	L	33.	H	L
10.	H	L	22.	H	L	34.	H	L
11.	H	L	23.	H	L	35.	H	L
12.	H	L	24.	H	L			

TEST 11

The items of this next test are made up of sequences of chords. What you are asked to decide is whether the first and last chords are the same or different.

If they are the same circle "S" on your answer sheet.

If they are different circle "D" on your answer sheet.

Here is practice item A:

S D

The first and last chords were the same, and you should have circled "S".

Here is practice item B:

S D

You should have circled "D", because the first and last chords were different.

Now try practice item C:

S D

You should have circled "S", because the first and last chords were the same.

Now turn the page, and get ready to try the test.

TEST 11

1.	S	D	16.	S	D
2.	S	D	17.	S	D
3.	S	D	18.	S	D
4.	S	D	19.	S	D
5.	S	D	20.	S	D
6.	S	D	21.	S	D
7.	S	D	22.	S	D
8.	S	D	23.	S	D
9.	S	D	24.	S	D
10.	S	D	25.	S	D
11.	S	D	26.	S	D
12.	S	D	27.	S	D
13.	S	D	28.	S	D
14.	S	D	29.	S	D
15.	S	D	30.	S	D

TEST 12

In this test we shall present sounds which keep a constant pitch, but vary in quality. Each item will consist of two notes, and you should decide whether they are the same or different.

Circle S if you think they are the same.

Circle D if you think they are different.

So now you are listening for a change in quality, or timbre.

Here is practice item A:

S D

Those were different, and so you should have circled "D".

Here is practice item B:

S D

Those were the same. You should have circled "S".

Practice item C:

S

Those were different. You should have circled "D".

In the test the number of every fifth item will be read to help you keep the place. BE SURE TO ANSWER EVERY ITEM. If you are not sure of the correct answer, use your best judgment.

Now turn the page, and try the test.

TEST 12

1.	S	D
2.	S	D
3.	S	D
4.	S	D
5.	S	D
6.	S	D
7.	S	D
8.	S	D
9.	S	D
10.	S	D
11.	S	D
12.	S	D
13.	S	D
14.	S	D
15.	S	D
16.	S	D
17.	S	D
18.	S	D
19.	S	D
20.	S	D
21.	S	D
22.	S	D
23.	S	D
24.	S	D
25.	S	D
26.	S	D
27.	S	D
28.	S	D
29.	S	D
30.	S	D
31.	S	D
32.	S	D
33.	S	D
34.	S	D
35.	S	D

TEST 13

The next part is a test of Tonal Memory. In each item you will hear a series of notes played twice.

In the second playing one note is changed. You are to decide which note is changed: the first, the second, the third and so on, and circle that number in your answer book.

There is always only one note in the second playing which is different from the corresponding note in the first playing. As you hear the notes in each set count them silently to yourself: one, two, three, and so on, so that you may identify the note that is changed.

There are thirty sets of notes in the test.

Here is practice item A:

1 2 3

The first note was changed, and you should have circled the "1".

Here is practice item B:

1 2 3 4

The second note was changed. You should have circled the "2".

The number of every fifth item will be read to help you keep the place.

Now turn the page, and get ready to try the test.

TEST 13

1.	1	2	3		
2.	1	2	3		
3.	1	2	3		
4.	1	2	3		
5.	1	2	3		
6.	1	2	3		
7.	1	2	3		
8.	1	2	3		
9.	1	2	3		
10.	1	2	3		
11.	1	2	3	4	
12.	1	2	3	4	
13.	1	2	3	4	
14.	1	2	3	4	
15.	1	2	3	4	
16.	1	2	3	4	
17.	1	2	3	4	
18.	1	2	3	4	
19.	1	2	3	4	
20.	1	2	3	4	
21.	1	2	3	4	5
22.	1	2	3	4	5
23.	1	2	3	4	5
24.	1	2	3	4	5
25.	1	2	3	4	5
26.	1	2	3	4	5
27.	1	2	3	4	5
28.	1	2	3	4	5
29.	1	2	3	4	5
30.	1	2	3	4	5

TEST 14

This test is much like the one in which you were asked to listen to the quality of the tone and to decide whether two notes were the same or different. However here the task has been made a little more difficult by having a five- or six-second interval between the two sounds.

You are to circle S if you think the timbres of the two sounds are the same,

D if you think the timbres of the two sounds are different.

Now try practice item A:

S D

Those were different, so you should have circled "D".

Here is practice item B:

S D

Those were the same, so you should have circled the "S".

BE SURE TO ANSWER EVERY ITEM. If you are not sure of the correct answer, use your best judgment.

Ready:

1.	S	D	13.	S	D	25.	S	D
2.	S	D	14.	S	D	26.	S	D
3.	S	D	15.	S	D	27.	S	D
4.	S	D	16.	S	D	28.	S	D
5.	S	D	17.	S	D	29.	S	D
6.	S	D	18.	S	D	30.	S	D
7.	S	D	19.	S	D	31.	S	D
8.	S	D	20.	S	D	32.	S	D
9.	S	D	21.	S	D	33.	S	D
10.	S	D	22.	S	D	34.	S	D
11.	S	D	23.	S	D	35.	S	D
12.	S	D	24.	S	D			

TEST 15.

This last is a memory test, pure and simple. A series of numbers will be read out, and then the word NOW. When the word NOW is read you should write down all the numbers in the exact order in which they were called out.

Do not write any numbers until the complete series has been read. The announcer will say NOW to indicate the series is finished and you may then begin writing immediately.

The series get progressively longer in the two halves of the test.

Here are two practice items:

Series A:

Series B:

Check your answers as they are read out:

Series A: you should have written ....

Series B: you should have written ....

Some of the series will be too long for you to remember all the numbers. If you do not remember some of them leave a blank space for them, and write down the numbers that you do remember.

Remember: write down all the numbers in the order that they were read, and do not write any down until you hear the word NOW.

Turn the page, and get ready to try the test.

TEST 15

1.

2.

3.

4.

5.

6.

7.

8.

9.

10.

11.

12.

13.

14.

15.

16.

17.

18.