A STUDY OF THE ABILITY OF RUSICIANS TO DFTBCT RELONIC AND KHOLONIC ERORS IN THE PBREORAMEE OF CHORAL

UUSIC MILIE MSPECTMO THE SCORE
by

Louis A. Hansen
A.B., Whessouri Valley College, 1937
B.M.g Wifsourl Valley College, 1938
M. H., Northmestem University, 1942

DISS
1955
Hansen
music
Submittod to the Department of Jusic Education and the Faculty of the Graduate School of the University of Kansas in partial fulfillment of the requirements For the degree of Doctor of Philosonhy.

Advisory Comnittee:

LIST OF TABLES ..... vii
I. THE NATURE, WEED AN PURPOSE OF THE STUDY ..... 1
Introduction ..... 1
Need and purpose of the study ..... 7
Statement of the problem ..... 12
II. REVIET OE RELATSD RESEARCH AND LITERATURE ..... 14
Musie achievement tests ..... 15
Tests of musical aptitudes ..... 25
Summary ..... 40
III. THE CONSTRUCHION OF THE TEST ..... 44
The rationale and general procedures ..... 44
Exploratory form of the test ..... 52
Reproduction of the test ..... 53
Method of marking emors ..... 54
Method of scoring ..... 55
Correction for guessing: general considerations ..... 55
The sample and results ..... 60
The preliminary test: part 1 ..... 63
Reproduction of the test fom ..... 64
Recording of the test ..... 64
Method of marking errors ..... 67
Wathods of scoring ..... 67
Questiomaire ..... 67
The sample ..... 68
Comparison of scoring methods ..... 68
Rellability ..... 74
Evidence of validity ..... 77
Item analysis ..... 77
Item difficulty ..... 78
Item diseriminating power ..... 80
Results of item analysis ..... 86
The preliminary teats part 2 ..... 87
Selection of additional excerpts and Introduction of errors ..... 87
Heproduction of the scores ..... 87
Recording of the test ..... 87
Harking and scoring ..... 88
Guestionnaire ..... 89
The sample ..... 89
Relfability ..... 90
Analysis of item difficulty ..... 90
Analysis of item discrimination ..... 90
Sumary ..... 92
The final form of the test ..... 92
Practical time limits ..... 93
Inolusion of non-discriminating items ..... 95
Order of ftems ..... 96
Representativeness of choral masic included in the final form ..... 97
Representativeness of chordal vocabulary ..... 98
Reproduction of the musical scores ..... 103
Recording of the test ..... 103
Test Instructions ..... 104
IV. THE DESIGN OF THE EXPERI期TLL STUDY ..... 105
Introduction ..... 105
The questionnaire ..... 106
The sample ..... 108
The eriterion group ..... 109
Statistical tools used ..... 111
V. RBSULTS AND INTERPRETATIONS ..... 115
Subjects classified according to amount of theory training ..... 115
Poaling of scores of subjects in each category of theory training ..... 118
Reliabillty of the test ..... 123
Comparisons between groups classified according to amount of theory training ..... 127
Validaty of the test ..... 129
Additional procedure for validation ..... 134
Comparisons of school marks in music courses with test achievement ..... 136
Reliability of the questionnaire data on school marks ..... 136
Comparisons in test achievement between subjects recelving "above average" and subjects recelv- ing "average or below" marks in music courses ..... 141
The relationship of selected factors to achievement on the test ..... 144
Age ..... 145
Bex ..... 145
Vozce part usualy sung in msemble ..... 148
Principal mediun of performance ..... 148
Piano experience ..... 154
Piano stucy ..... 157
Voiee stuay ..... 157
Voice experience ..... 160
Principal instrument, other than piano, of Instrunental majors ..... 163
Arsunt of strudy of principal instrument by Instrumental majors ..... 167
Lencth of experience on major instrument of instrumental majors ..... 167
Piano sight reading abiluty ..... 172
Piano improvising ability ..... 172
Choral directing experience ..... 175
Instrumental directing experience ..... 177
The nethod used predominantly in Ilstening for accuracy of harmonic effect ..... 177
Anslysis or recognition of chord function ..... 183
Sumary of resulta ..... 185
Addenda ..... 189
VI. SUMAARY, CONGLUSIONS, AND RECOMATSDATTOMS ..... 191
Surmary ..... 191
Conclusions ..... 197
Recomendations ..... 199
Concluding statement ..... 201
CRAPTEE ..... PACE
BIBLTOGRAPHY ..... 203
APPEMDIX A Original Data ..... 209
APFBNDIX B Exploratory Test ..... 232
APPETDIX C Prelininary Test, Part 1, Relfability Data, and Item Analysds Data ..... 249
APPENDIX D Preliminary Test, Part 2, Rellability Datas and Ttem Analysis Data ..... 270
APPENDIX E FInal Test and Questionaire, Item Seleotion Data, Analysio of 鲑ical Quality and Function of Test Iters, and Reliability Data ..... 278
APPRMIX $F$ Letter to Heads of sinsic Departmente ..... 296

## LIST OF TABLES

TABLE
page

1. Classifioation on 1,000 Applicants According to
Reading Ability 48
2. Classification of 582 Applicants Tho Indicated Some
Training on tho College or Conservatory Level. . . 49
3. Ranks of Uncorrected Scorea and Scores Oorrected for
Guessing or Chance Sutcess, Using a Gross Correction
of -5 Ior Each Incorrect Response Exploratory
Test
4. Ranks of scores Obtained by Nethod 1 (Weighted) and by
Method 2 (Unweighted) Part 1 of Preliminary Test. . 69
5. Ranks of Uncorrected Scores (Obtained by Hethod 1), and
Scores Corrected for Guossing or Chance Success,
Using a Gross Correction of -5 for Each Incorrect
Response Part 1 of Prelunary Test . . . . . .
6. Item Difficuity of the Praliminary Test, Part I*
Huber of Items Found for Nach Devel of Diffim
culty Represented
7. Iten Difficulty of the Proliminary Test, Part 2 : Number of Items Found for Each Level of Diffil-

8. Choral 期sic Represented in the Final Test Form *... 99
9. Sumary Data of Analyses of the Quality and Function of the Chords and Intervals of the Final Test, According to Their Notation, into Which Performance trrors were Introduced102
10. Schools Participating in the Study, and Number of
Test Rosults Received from Each ..... 110
11. Members of the Criterion Group, and Qualifications for Hembership According to the General Criteria ..... 112
12. Subjects Classifted According to Amount of Theory Training ..... 217
13. Analygis of Horogeneity of Distribution of Scores of Subjects Having One Year of Theory Training: Schools $A, B_{\text {, }}$ D, and $E$ ..... 119
14. Analyois of Homogeneity of Distribution of Scores of
Subjects Heving Two Years of Theory Training*
Schools $A_{i} \mathrm{H}_{\mathrm{y}}$ and $\mathrm{D} \ldots \ldots \ldots 121$
15. Analyess of Homogenelay of Distribution of Scores of
Subjects Having Three Years on Hore of Theory Traintrig: Schools $A, B, O$, and $T$ ..... 122
16. Proportion of Gases Contributed by Groupe to a Stratified Sample of Fifty Cases ..... 124
17. Individual Cases Erom Groups in a stratifiled handom Sample of Tifty for Uee in-Calculating Test Rellabiluty ..... 126
18. Comparisons Betreen Groups Classified According to Anown of Theoxy Training: tm and d-Tests ..... 128
19. Analysis of Fomogenelty of Distributhon of Scores of Groups Having Tho Tears of Theory, and Three or Woxe Years ..... 130
20. Distribution of Test Soores and Ranke for tusicians Grouped According to Amount of Theory Training, and for the Oriterion Group ..... 132
21. Comparisons Between the Critorion Group and Each Group Classislied According to Amount of Theory Training ..... 135
22. Comparisons in Test Achlevement According to Marks in Husic Theory Courses as Reported by Subjects and According to Actual School Marks ..... 139
23. Comparisons in Test Achievement Between Subjects Receiving "Above Average" and Subjects Roceiv- ing "Average or Belom" Marke in 䱣usic Theory  ..... 142
24. Analysis of Homogeneity of Distribution af Scores According to Age ..... 146
25. Analysis of Domogeneity of Distribution of Soores According to Sex ..... 147
26. Analysis of Homogeneity of Distribution of scores According to the Part Sung ..... 149
27. Analysis of Homogenalty of Distribution of Scores According to Principal Mediun of Performance ..... 150
28. Analysis of Momogeneaty of Distribution of Scores of Piano Majors and Voice Majors ..... 152
29. Analysis of Homogeneity of Distribution of Scores of PLano Majors and All Others ..... 153
30. Analysis of Homogeneity of Distribution of Scores According to Hunber of Years of Piano Playing Experience ..... 155
31. Analysis of Homogenaity of Diatribution of Scores of Subjects Hiaving up to Five Years of Plano Playing Experience ..... 156
32. Analysis of Honogenelty of Distribution of Scores of Subjects Raving up to Hive Years of Plano Stuxy ..... 158
33. Comparisons Between the Distribution of Pooled Scores of Subjects Having up to Five Tears and that of Subjects Having Six or More Years of Plano Study ..... 159
34. Analysis of Homageneity of Distribution of Scores of Voice Majors According to Number of Tears of Voice Study ..... 161
35. Analysid of Homogeneity of Distribution of Scores of Voice 期jors According to Years of Singing taxperience ..... 162
36. Analysis of Honogeneity of Distribution of Scores of Voice ${ }_{\text {laju }}$ jors According to Anount of Piano Experience. . ..... 164
37. Analysis of Homogeneity of Distribution of Scores of Voice tajors According to Broad Categoriea of Piano Experiance ..... 165
38. Analysis of Homogeneity of Distribution of Scorea of. String, Brass, and woodvind kajors ..... 166
39. Analysis of flomogeneity of Distribution of Scores of Instrumental Hajors According to Amount of Study on Their Principal Instrument ..... 168
40. Analysis of flomogerieity of Distribution of Soores of Instrumental kajors According to Amount of Experience on Their Principal Instruent ..... 169
4i. Analysts of Fomogenelty of olstribution of Pooled Scores of Instrunental Wajors Classified Accord- ing to Amount of Plano Experience ..... 171.
41. Compardson of Test Achievement of Subjects Able, and Subjects Mot Able, to Play Simple Piano Rusic at Sight ..... 173
42. Comparisons Between Test Scores of Subjects Able, and Unable, to Improvise Plano Accompaniments for Simple Melodies ..... 174
43. Analysis of Homogeneity of Distribution of Scores According to Humber of Years of Choral Directiog Evperience ..... 176
44. Analysle of Homogenetty of Distribution of Scores According to Number of Years of Instmmental. Direating Expertence. ..... 178
45. Analysis of Homogeneity of Distribution of Scores According to the wethod Used Predominantly in Listening for Accuracy of Harmonic effect ..... 180
46. Comparison of Kethod 1 und Method 2 os Listening for Accuracy of Hamonte Effect ..... 181
47. Corparison of Hothod 2 and Mothod 4 of Listening. for Accuracy of Harmonic Bifect ..... 182
48. Analysis of Honogeneity of Distribution of Scores According to thatent of Analyais or Recognition of Chord Function ..... 184
49. Data for Musioians in School A ..... 210
50. Data for 值icicians in School B ..... 217
51. Data for Pasicians in Sohool O ..... 221
52. Data for iasicians in School. $D$ ..... 223
53. Data Sor Musicians in School E ..... 226
54. Data for Musicians in School F ..... 228
TABLE ..... PAGE
55. Data for 学usicians in School $G$ ..... 229
56. Data for Wusicians In School H ..... 230
57. Data for 猜usicians in Criterion Oroup ..... 231
58. Preliminary Test, Part 1: Scores Obtained for Split. Halves of Test, Using Alternate Itemg, for Use in Distinating Test Reliabiluty ..... 264
59. Indices of Difficulty and Diseriminating Power of Individual Itemst Prelininary Test, Part 1 ..... 265
60. Preliminary Test, Part 2: Scores Obtented for Split Halves of Test, Using Alternate Items, for Use in Estimating Test Heliability ..... 275
61. Indices of Difficulty and Discriminating Power of Individual Items: Prelininary Test, Part 2 ..... 276
62. Excerpts and Itens Selected for Ver in Final Form ofTest, Including Indices of Diseriminating Powerand Level of Difficulty of Items287
63. Final Tost Form Analysis of the Quality and Function of the Chordes and of the Structure of the Belodic and Hamonto Intervals, According to Their Nota- tion, into Whith Performance Errors Were Introduced ..... 291
64. Final Test: Scores Obtained for Split Helves of Test. Using Alternate Itern, for Use In Estinating Hest Rellability ..... 295

CIAPTER I

THE MATURE, REED AND PURPOSE OF THE STUNY

## Introduction

The ability to read and hoar music accurately is inportant for musicians. Such ability has special gignificance for those who train and direct musical eroups. Teachers of music in the schools obviously should have an aciequately developec techncque in reading, and an ear trained to meet the practical moolems encountered in chorel and orchestral conductm inc.

Jusic teachers are obligated by the nature of their profession to dovelop a high degree of still in reading the language of music. They mast examine quantities of material for choral and instrmental roups; the practical value of repid and accurate hearing is clear. Facing a chorus or orchestra, the conductor mact surpass the group in his ablity to grasp the minted page. A thonoughly developed aural sense is equally importent, for the conducior must instantlyr detect rivthenic and tonal errors.

This study was particuarly concermed with the ability ${ }^{2}$ of musicians to read and hear the nelodic and hamonic intorvals and choras, as inportant eloments of music, because the music cirector must be continually concerned with having the dotes perfomed as mritten.

Roading musical notation is a process of obtaining meaning fron symole, and presupposes the aidility to thini: in musical lenguage, just

[^0]as in the reading of English it is necessary to get meaning from words, and presupposes the ability to think in that language. Wotes and chords have a definite meaning only to the reader who has a definite knowledge of the tones and hamonies which these symbols represent.

There are important and subile relations betreen the hoaring and readin processes, for

The porson who hears mentally what he sees reads best. It is possible, thmough various mechantoul methods, to read quite rapidly without nuch consciousness of the aural affect. This is not a very intelligent method of approach and is a poor basis for musicianship. Rost nusic oducators mill admit that what we are aiming for is intelligent reading of music at sight. Another is eccurate mental hearing of the printed page. 1

Reading and mental ineging of aural effect are closely related and neariy simpltaneous processes. Just as the skinled reader of printed words can imagine how they hill sotid, so the sicilled musictan can inagine the sound of the musical symbols. Agnsm, ${ }^{2}$ in a stady of the power to inage and think in tons, fowd that misicians almays roported the use of auditory imagery. And more than half of the mucicians who served as subm jects attested that auditory inagery could be as vivid as actual toral perception, She also foundg fron a study of thein lattors and autobiom eraphes, that the eroat composers lived in a tonal world in which actual hearing frequentily was of less importance than mental hearing. They

[^1]constantly emphasized aural inagery in their advice concerning musical training. ${ }^{1}$

Proquently ono bears a musician attost to his skill in hearing music mentally from its notation.

I renember a gathering, mainly of literary peopla, in which I was fakon to task for fmplyine quite casually that the sound of an orcheatral score, and the possible interpretations of that score, could ve quite vivid to me through reading the notes, oven is I had neither heard the score perfomed nor played the notes rayself. ${ }^{2}$

Such highly doveloped powers of auditory inagery are expocted of slilled musicians. "Ary real musician is able to pick up a score and by nerely reacins the notes, hear the music accurately." 3

The accurate reading and hearing of music as an important aspect of musicianship has long been recognized. Evidence of such recognition is tho inclusion of courses in the curricula of nusic schools, generelly Iisted mder the tithe of theory, which have as a major purpose the development of this ability.

At the present time increasing emphasis in theory courses is beine placed upon the aural exporiencing and smaging of music in conjunction with its written aspocts, for after all, music is a tonal ant and is prinamily to be heard. The symbols of mesic aro only a means to that end.

[^2]A recent tertbook points out that "meic is heard, not seen, and so one's enalysis should be primarily aural, not visual. This means that music notation whon seen should be heard mentally * . .1

Bven for the ansical layman who may have littile ability to interm pret complicated scores, the value of thoory stady lies in the finer discrintination of anral offects.

The study of theory assists in maleng it possible to hear more. corpletely 271 details of the tonal complex that charactorizes our music. The Jistener traned in theory is more amere oin the details of myinas, melodic line, harminic content and musical structure thon the non-trained Iistener. ${ }^{2}$

In the area of gerformance, the power of aural discrimination sharpened by theoretical tratning has significant application.

A lmowledge of the functions of tones, resulting from a straty of theory, influences the pexAomer's rendition and leads to more effeotive perfomance. For exmple, middle $G$ as the thitrd of an Aminor chord performs a different function from the same pitch as the leading tone in the key of $D$ flat najor. Wifh an awareness of this functional difference, the singer ox preformer on an inm strument where pitch can be varied will alter his rendstion of the pitch to make it perfom its function betior. $G$ as the leading tone may be rendered slightly sharp; as the minor third, a fert viorations 2lat. Such slight variation enhances the function of that pitch within the tone complex and nates the rendition nore effective. Wary such variations resulting from the perionmer's knowledge of the relationship of tones, choris, and elements of the structure have their influence on the total effect of the performance and make the difference between an artistic and an orcinary rondition. 3
$I_{\text {Howard Ansley sumby and Hatrin Jonn Siringham. Creative Harmony }}$ and Wusicianship. Prenticemall, Inc., Hew Yorit, 1951. P. 95.

ZHazol Wohavec Worean (ed.) Basic Education Source Bool. Khuric Elucatore Hational Conference, hicagos IT. $194 . \mathrm{F}$. IIS.
${ }^{3}$ Tbia. P. 119.

For the music director the ability to read the score and understand its complexities is ereacly aided by the study of theory.

The score reader well versed in music theory can hear the score intemally better than the non-trained reacier. Since his inner inage of the tonal complox is fuller and more accurate, he will denend moxe from his performing eroups. The stendy of theory thus sharpens the ear of the prospective conketon.?

The inportance of aural inagery in its relation to score reading and the nenner of its doveloment through training in thoory has been sumburized as follows:

The aim today is to develop auditoxy imacery in the pupil so thet when he thinks "tonic chord" he is not merely seeing a chord on the staif but is heariag a comblnation of tones consisting basically of a major third vith a minor third over it. Sinularlyg mhen looking at the hamonized melody that he has mitten in his blanic book, ho will hear what he soes, just as though he were. playing it at the piano. And in the end he will cone to the point where my musical score which falle under his eye will arouse auditory inasery, so that he will actually hear the masic with his "inmer ear" as he looks at the notes. Not everyone will have the power to do all this of course, but overyone has emough talent to do at least sone of it. Else he outht not to be taring amsic theory comse. ${ }^{2}$ (Italios in the original.)

Basic theory wrik has ofton beon conparimentalized into separate carmbaning and hamony courses but the trend is toward their integration because in both, the nental hearing of tones and hamonies is of prime imporbance. The ear training aspects of theory study are, of course, specinically devoted to developing powers of aural discrinination and inagery. "Ho schone of ear truining can be considered adequate which

Izorgan. $\overline{2} .119$.
${ }^{2}$ Karl \%. Celmitens. TWudanental Principles of Theory Teachinge" MTM Proceedinge, 34 (1939) 64-71.
does not make constant provision for developing musical imagery. ${ }^{1}$
Practically overy theory tesoher bears mitness to the necessity of developing this skill. One such testimony will suffice: "The mosit vital part of the whole subject of theory teaching is ear training. It is inw possible to overstate the importance which the training of the ear should occupy in the minds of both teacher and pupil. ${ }^{2}$

Liserise in specific courses in harmony tho aural aspects are coning to be racognized as of primary tuportance, rather than the manipulation of symbols. The value of whiting hamony unaccompanied by aural imacery has been questioned.

A poll of practice shows that nore time in theory classes is given to the written aspocts of thcory (Sour-voice hamondzation of melodies, for example) then to any other one activity. It has been questioned whether the usual written activities contribute vitally to the reel onds of theory instruction. It is belleved that whitiny by rule uneccompanied by innor consciousness of the sound of the product is quite useless in developing a desirable musical goal. 3

In a study of hon 624 selected choral leaders actually used knom-
ledge of hamony and skilla involving harmony, it was found that the aural aspects of hamony were of yaramornt inportance.

Not oniy are the activities involving the hearing of harsony rated -as most used, bat in their very nature they are crucial to the choral leader's work. Silent study of the hamany of choral mubers, although not an absolute ossential, is an efficient way
$I_{\text {Janes }}$ L. Zursell and Mabelle GIenn. The Poycholow of School Masic Teaching- Silver Eurdoit 00., Hew York, 1936. P. 173.
$2_{\text {Angela Diller. First }}$ Theary Book. G. Schimor, Inc., Hew York, 1921. P. iii.
$3_{\text {Morgan. P. }} 120$.
of preparing monders and of selecting nubers for use. Waking sure that the chords rendered fur rehearsal are the same as those of the soore is an absolute essential of rehearsal procedure. The high cating of usefulness given these activities, and their essentiol nature, engect that the success or failure of the hamony training program for choral leaders must be determined by the exient to which these amral skille are attained. No matter what other skills may be developed in the harmony prorreng, if the choral leader cannot hear hamony from its romesentation in the score, the program of tratning has not been functional. Wajor attention in the training program shoula be given to this aspect of the subject; accoram plishnent should be measured in temis of the cxitent to which the student can hear what be geos. ${ }^{1}$

Heed and Purpose of the study

The prooeding discusbian hes shown that rapid and accurate score reading ability is necessary for masicians, espectajly for those wose activities demand its use. Score reading inrolves the zeadingg undergtanding, and mental hearing of the msical notation. The actuel hearing of the music confims, mocifies, or disagrees with the mental impressions produced by tho notation. It is apparent that eye and ear must be highly coordinated, so that the function of one becomes closely related to that of the otiner. Fron a practical viewpoint, the process appears to become one integreted sunction.

The ability to read and undorstand masical notation has an inportant use in the rohearsal situation. It is ossential that the tones and hamomies, as important structural elements of music, be heard from thein representation in the score. The purpose of correctly hearing the tones

[^3]and hamonies id to elininete tiscrepacies betmen their notation and their performace, enc thas to secure the correet perfomamee of the music. Skill in reading mathearlig tones and chords would tharefore be ovidencen by sitill in dstecting melotic and hamonic orrors in the perm formance of ausle while reading the soore.

Weiclanship is oretranily undergtood to include the ability to read a musical score, Meicianc who clain to be able to reac incepencentIV, and certainly those who need or are expeoted to use such ability in their masical activitios, should show some degree of proficiency on a test which measures the extent of sinll in detocting disorepancies botween the notation and the periommae of tones and hamonien. one choral director recomends that "every sehool sumbing dezres in music should refuse to equduate my person wo lucks the fundamental mubicianship regnired to pass a simple score reading test of this sort." ${ }^{1}$

But the actual exent of score rondine ability as indicated by such a test has not been lnow. In vier of the value of score reading ability, particularly in such a functional stituation as a rehearsal where it mast be applied to a pracitical mugical situation, some incication of the actual achievement of musicions would be useful in eotting up tranm ine progrome for musiciane, on this tertative basia, desirablo standards of achievement could eventually be established.

From the results of a test of this type, siter its use in an oxperinental situbtion, way be inferred the contribation of theory traning

Itara G. Hogeard. Inproving wusic Reading in the Choral Rehearsal. The aubhor, Hew Yoris, 1947. E. 24.
to skith in reacing and homine mutics bocause tho rocogntion of melodic and hamonic errore snolven a monleage of both the wristen and aural aspects of jutervals, seales end chords, when whok theory trainone is apocificaly concemad.

Furthemore, tit has beea enorm that knomlocge of harrony and shills involvine hamory have their groatest nge and nost prectical appication, at least for choral directors, in the gaining of nental concepte of nusic from the silent atudy of the score, and in the comparison of the sound of hamony with its notation during the rehearsal. It seems reasonable to judge activittes in theory twaning by their contribution to musical coals. hecurate score poadime is cortangy a logitinate goni gince a constatit activity of mustic directors is to moke wure thet the mucic is pexpomer as notated.

The actual contribution made by theory tramine to the akill of correlating eye and ear in a score reading test; such as natertaken in this study, hes not been known. Some knomjecige of tio relationshin of thoory braining to accomplishent on a test, which utilizes such training in a very useful sithation, would be of considerable value to teacherg of theory in dotemining the methoa and content of theory courses.

A test of ginill in detecting disoropancies between the notation and perfomance of the tonal content of music would also help to fill a need for an acherrement test in this area. On the college level the training of musicians for teaching positions and other directing responsibilities
$1_{\text {Burns, op. cit. }}$




 in that it is rehathe to Jone range ablectives - bo bohavory and sicills found waent an the iob.

Some krowlecte of the gronth of such a complow akill as soore

 training in theory may anter into the dovolopont of boro reading 3kil.


 skill.

In view of thse constamations, this study wes undertaken to investigato the seore reacing ability of musiolans, gpecifically the abjlity to read and hear the melodic and harmonic olenento of nasic, as ovidenoz by the Aotoction of melohe mad hamonic armors in the performance of maic timila inspecting the sore the investigation was contucted among mojcians ether prepering for; or alresiy actavo in a musical carecr, namsly undergraduate and graduate musie majors in colleges and universities.

[^4]The study also was to be concerned with the rolationship of cortain factors or characteristios to this ability. The factore to ve selected would be those which might reasonably be expected to have some relation to tims ability, and which coule feasibly be included in the study. One obvious factor would de the anount of training in theory. This would involve the dotermination of levels of achievement in theory. Anong other factors to be invostigated were: the principal medium of performance, directing experience, keyboard facility, anoum of study and experience on the principal medim of perfomance, age, and sex. It would be useful also to determino which meblod wes used predominanty in tistening for accuracy of hamonic effect, and the extent of analysis or recognition of chord function, and their relation to sicill in score reading.

The test of reading and hearing the melodic and hermonic eloments Hould appear to be bast confined to the choral modiung for the following roasons: (I) The voice is a natural instrment, and it is probable that every musician has had some emperience in choral singirg and in tistening to chorel music. It seens not as probable that every musician has had a comparable amount of instmmental experience. (2) The vocal nediun is used to a great extent in meny pheses of theory training, such as sightm singing and dictation. (3) Althowh piano conld be used as the mediva of performence, sinco it is widely used not only in theory training but is also the mosi commonty used neans of gaining concopis of the hamony, in a choral situation it is tikely that singers create more mistales in tones and hamonies than does trie pianist. Desides, the director would probably be concomirating nainly upon the efforts of his singers.
(4) In a choral rehearsal situetion aural skill has been found to be of prime functional itportance in correlating the sounds with the notation of harmonies. There is, however, no ovidence that this would not also be true in an instrmental rehearsal situation. (5) A prine consideration in studying the skill of reading and hearing the score would be to crente a functional situation approaching, so iar as possible, a rehearsal situm ation in which exrors in performance could be introduced logically, but in such a maner as to tost simultaneously aural and visual sicill, and not merely musical sensitivity. It was believed that the choral medium Fould be more adaptable for this purpose. These considerations have been amplified in Chapter III.

Since the test was to be in a choral medium, and therefore would be of prinary intesest to choral directors, an effort was made to have tino musical extaples be fairily representative of the choral literature which the average high school, college, or church choir might use.

Decause the test would measure a skill emphasized in theory traininf, and therefore would be of interest to theory instructors, an ettompt ras made to sample at leest the comonly used chords, although it was realized that the contextual masical factors could affect the way an itern Iunctions.

## Statonont of the Problen

The rrorpose of the study was to mesure, among musicians preparing for musical careers, the extent of the ability to detect molodic and harm monic discropancies betwoen the notation and the performance of choral music, to detemane some levels of achievenent in this abiltity according
to unsic theory tratuins siatus; and to investigate the relationship of selocted factors to the ability.

To achieve this parpose, it was necessary to construct a valid and relieble insimment of neasurement. This becone the finst objective of the study. In the construction of the test two subsidiany objectives were sought: (1) that the test contain a pepresentative smple of musio which an average choir magt ase; (2) that the chords in which errors of perfomance occurred be a representative sample of comonity used chords.

In order to detemine levels of achevement, according to nusic theory troining status, in the ability neabured by the instrunent, and to determine the relationship of other selected factors to achievenent, a questionnaire to obtain data had to be devised. Its fom and content, and the analyais and interprocation of the obtained data, constituted the second objective of the study.

The adninistration of the instrument to a representative sample of the growps for which it was intended was necessary in order to achieve those objectives.

## OHATE $2 T$

## REVIEN OF RELATED RESEAFOH AND LTTERATURE

Sowe of the literature conceming the nature and ingortance of
 zad and hour mate as dereloped througn tratning and exporiencey she ta themefore bect measurod by nome type of achlevement tasty

In order to ahow the acturl need for construmtion of an antitional meabure, many of the curment maio achievoment testa are reriowed and evoluated. Testa of musicel aptitade, althong desimed to messure natite cbilithes os capacities, are also reviewod briofly for the further informetion that they being to the general area of maja testing philosophy and procedures

A manber of studies inge deat with other argas of musical leamm ing. Nost numerous ere the investigations relating to the reading of masical notetion. These hro dealt, anoze othon thing, with the amatysis of eye movesents of trained musiciansy amatewes and paxsons without majow cal traning; the effect of tho printed pase mpon masic reading; the ofSect of the grouptig of melodic pattems mud vertional extent of chocdss the effect of spacing symbols according to rivthmic ratio; the use of the tachistoscope in improving reacing ability by training the eye to perceive notes rapidy in priterms; and malyses of general characteristics of good readers. These studieg, though usciva, are not revieved here, since they are insufficiently rolated to the prapose of the present study a sum mary of the inportant findings of these studies, as well as of other research in the general sleld of musieal learnint, has been prepared by

Henorickson and Stratengyer. ${ }^{\text {I }}$

## Music Achtevement Tests

Thers ara a mabor of taste dostgned to meesure the results of truthing ond exporienco which may, therefore, be designated as achievem

 spread uee, memures howledee or school nusic from grades 4 to 12.
 vary find five to tumaty-nve atems.

1. Frovieledea of musioal symbols and tamm.
2. Recogrition of syllable nones from notation.
3. Detaction of pitch arrors in the notation of a lamiliar nelody.
4. Detection of tince emors in the notation of a fantiliar melody.
5. Rncwletge of pition or lefter names of bass and trevie clef.
6. innomedec of tine signatures.
7. Krowleage of key sientatures.
8. Knowledge of note values.
9. Whutlucge on róst values.
10. Recogntion of familian melocies from notation.

The relianility of the test is reported in the manal of directions as .97 by tho split-half methoc. The veltdity is based unon a standard course of music study recomencled by the Itisic Supervisoxs Hational Conference and upon the courses of study of soveral outstanding public
$7_{\text {Gordon }}$ Hendriclsson and Clara G. Surateneyer. Mugic Education." Fncyclopnedia of Educatiomal Resempon, ed. Waiter 3. Homroe, The Hacmilian Company, hem York, 1050, 701-771.

Jacob Kralwasser and G. If. Ruch Kwalvasser-Ruch Test of husical Accomplishent. Bureu of Bducabional Research, Eniversity oi Iowes Iowa CJty Iowa 1924.
achool syotems．The noms provided are based upon 5，hi4 test rosults．
 tests theoretical knowledget note and rest values，tirie signatures， pitch and syllable nemes，expression marks，repest signe，major and minor key signatures，and natural and harmonit mitror scales．Part Beasures the ability to supbly，from aural dictation，corect syllable nemes in twelve exerciees，to write time vignatures and supply measure bars for four given relodies，to detect pitel and tine errors in notationg，and to Write notos on the steff from diotation．The eaphesis given to the writu－ Ing of music has been oritioized by masic supervicors ac a waste of the pupil＇s time．

The Eutchinson 解基 Tests ${ }^{2}$ are similar to tho second part of the preceding test，In thich femiliar melodies are to be recoenized by silent stuafy．Trenty－four phrases Erom famhar songe，divided Into six groups， are to be identified from the list of eight song titles printed below each eroup．The teets are intended for grades 7 through 12.

The Beach 誨sic Tast 3 has eleven parts，for the Items of whath correct answors are to be selected from three or more printed choices．

[^5]1. Inowledge of music eymbols (by recognition from the notation).
2. Recognition of measure (recognizing the number of beats in measures played to the subjects).
3. Recognition of tone direction and ginilarity (indicating direction and repetition in melodio phrases played to the subjects).
4. Pitch discrimination (deciding in various gituations which tone is higher or lover, highest or lowest).
5. Application of gillable names to different pitches plagred, without notation.
6. Tine values (indicatine the time-placoment of designated notes or note patterms).
7. lusical terns and symbols.
8. Correction of errors in the notation of a heard melody, insertion of time signatures, and identification of familian melodies from their notation.
9. Writing of syllable and pitch names (according to notation).
10. Representation of pitches (mating dom in notation the pitches of various notes heard).
11. Jatching composers and artists with brie. characterizabions of them.

The items were selected on the basis of teacher opinion. The test \#as validatod aganst toacher rankings on knowledge of musical fundamentals, and on general musicianshtp. Correlation of the test scores of 535 pupils from a four-state area with the rancings fielded a coefficient of .67-.02. The correlation of tert scores with the rankings of college music students rated on sightmsinging ability, kowledgo of musical fundomentels, ance general musical ability was $.87^{+4}$. 02 . The reliability of the test was determined by comparing the scores of several combined parts. The average split-half coeficients obtained were . 86 for college students and . 83 for high scinool students. Percentile grade noms are given for erades 3 through 8, and for high school as a unit.

The Providence Irventory Test in Tusic $^{1}$ is confined lergely to kowledge of musical symbols. The ten parts include naming notes, placine "do", naming note values, naming key signatures, naming measure gignatures, namins rest vazues, neming syllables of the treble staff, naming familiar melocites from notation, naning syllables of the bass staff, and naming other symbols. The noms are based on the test results of 15,000 Providence school children.

The Gildersleeve Husic Achtevement Mest ${ }^{2}$ is in four parts, with an additional section for the gaining of backrround information.

1. Komledge of the way musicel instruments are played.
2. Knomledge and use of mosical symbols, inoluding placing the keynote of given signatures, writing syllables or soale numbers belom criven notes, miting pitch nanes, and placing bar lines according to time signatures.
3. Knorledge of types of compositions, fanus composers and musical verms.
4. Satching the notation of familiar melodies with their titles.

The test itcns were selected by fifteen advanced music students. The reliability of the test was reported ${ }^{3}$ as $.905^{\circ}$. Moms are based on the scores of 3700 puoils in grades 4 through 8, from trenty-iour school systemas.

[^6]The Knuth Achicvement Test in Music has as itw objective the moasurement of the avility to recognize masic from its notation, ancis terned a test of sirht-reading ability. Tho forms are provided for each of three divistons, grades $3-4,5-6$, and $7-1.2$. The test contains forty progressively difficuli itens. Tach zten consists of a four neasure phrase minch is played. The first two neasures are given in notation, with four altornative notations given for the last two measures from which the subject is to select the one mich is playod.

This test goee beyord the mere knowledge of notational symbols in that it deals with their actual perception. Validity was establishod by an analysis of a numer of public school music series texts, by pooled oapert judgmonts on each of the test items, and by an experimental tryout using over 4,000 cests mich secured the percentage of then successes at each grade level. For grades 7-12 the comparable forme reliability was found to be . 84 with 1109 test scores. With extering teachers college students, in a tost of the entine range of the tost, the two forms yielded a coefficient of reliability of. 953.

The recent Diegnosige Tests of Achievement in Music ${ }^{2}$ are similar to the earlier achievement tests which sinnly measure kowlecige of symbols and tems. The ten tests include the theory and notation usually presented An grades 4 to 6 , as determined by an analysis of current mance series
1.flliam E. Kuth. Knuth nchievenent Rests in Kusic. Bducationel

${ }^{2}$ Lela 3 . Kotick and T. T. Torgerson. Diagnositic Tests of Achieve$\frac{\text { ment }}{1950 \text { in Music; Mamal of Directions. Califormie-Test Bureau, Los Angeles, }}$
textiooks. The tro form of the tost may be used from grade 4 through
 Wine of the ten parts are prinded on the anwer sheet; Part 10 is printed on a separato meot. The tests are: (1) Diatonic syllablo names, (2) Oromatic syllable names, (3) Number manes, (4) Time signatures, (5) Major and infror keye, (6) Note anc rest values, (7) Letter names (treble clef), (3) Signe and sybbolss (9) Koy names, (10) Sonc recogrition from notation.

Wh norma are provided. the anthors believe that they would have little meaning because of current diversity of music curricula, lack of uniformity in the grecie placemont of topies, and varying emphasis upon syotematic theory imsining and steht-singing. They recomend that tio togts be adated to loeal stumetions, with rew scores being convorted into percentace scores for sach test and acceptable standards being set where the testr agree with the curriculum.

The reliabilities reported for each test are based on more than 200 cases from eredes L through 10, with the Fuder-Richardson fomma aphied, and are adequete mith only a fem excentions. Meliabilities basci on total test scores for each grace rance fron . 39 wo . 98.

An example of a test wich sampes acquirod information about music and composers, inke parts of the krabrasser-much and aorgersonWemestock torts, is tho Fralmasser Tost of Musticel Infomation and Apomeciation 1 It is invended for use in mistc apyreciation classes.

IJacob Krralwasser. Hwalwasser Test of Musical Information and Appreciation. University of Iowa, Iowa City, 1927.

There averina dubtants:

1. Orasificetion of surtista.
2. Nationelity of composers.
3. Composers of ingous moviss.
4. Classification of composers by types of composition.
5. Generat historical and biographien moviedge.
6. Procuction of tones of ozehestral instiments.
7. Clawsificaion of orchentrel instmanents.
8. General kowledge of inetmontation.
9. Knowledeg or music structure and form.

Several tests are concemed mith the direct measurenent of sightsinging abtity. The M11brand ghemsinging Teat for Grades in to 6 is an individual, cral test containing six songe Their tonal content Includes upwan and domward scale pogressions end tonic chord progressions. The simple neter signetmes are mployed, and the rifthuc range is confined to quattor, helr and eighth notes, with the quarter note always the unit of beat. The test procedure is to have two pmil stucy the notetion for a faw minutes and thon sing the sorgs winont assistonce. The praninex, ustak prescribed eligns, is to mark on a copg of the sones such orrones as notes wrongly pitchec, tranapositions, fiatitug or sharpm ing, axtra ur omitted notes, and tine emrors.

The test is therough but the grading systen by its nature is aifficult and timecoasuming A mamal of directions and noms are availablo.

[^7]Sinilar tosts of sighe-singing ability have been devised by Gews ${ }^{1}$ Salisoury and Sxith, ${ }^{2}$ and 0ttersteln and Wosher. ${ }^{3}$

The difficulties of administering and scoming individual teste of slehtminging ablity led Mosher to atheaph bo devise some grony tests ${ }^{4}$ which, while not meacurine slithosinging abilsty directly, would be a significan inder of it. Mosher anazzed school texis and teachers' manvals and found that the subject matexiel which supposediy contributed to achievement in sight-singint conld be divided into seven phases. For each of these he deviced grow tests:

1. Knowledge of musical syniole, marics of expressiong and general music information.
2. Rocognition of ceales, chords, and jntervals.
3. Nomledge of measure and nota dumation values.
4. Ability to identify meli-kom meloaks when read silently.
5. Ability to mitce tonal figumes or putiems from hearine then played on the piano.
6. Ability to write rathmic pattema from hearing then played in monotone ors the piano.
7. Ability to write nelodies from dictation.

The degree of relationship betreen the perfomence on an individual test of sightringing ability also devised by 10 gher and these group teats prove investigatod. The individuel test was scored by suming the maber

[^8]of meagures which had no tonal errors and those which hed no rintimic errons.

The group test most indieative of sight-bincing ability was found to be test No. 4 rith a correlution of 62 , the ability to wite tonal figures from hearing them. The multiple corvelation between total test scorcs and the individual tests wae .60. Wosher concluded, in part, that in it be trie that reading menc is cependent apon the power to think macic, "ability to vinto from dictation the severnl tonal patterns is uncoubtedly a neasure of one's power to think til the tonal languge in terme of relutionsilip of pitch. ${ }^{2}$

There are many achevemont cests of various types designed for stucents apecialiaing in music in high school and collegeg but few have boon stindardized or reported in the 1 tierature. The tests by witght are an erangle of those few. The firgt test neasures the periomence on the piano of tro selections from the stucent's repartoive, one beine chosen by the oxaminer and one by the student, pano sightroading ability of school music matorial, the ability to ainc back a four-acasure molody, and the ability to mite the notation of a melody fiven on the piano. The second test measures the abitity to write amsic from dictation, and the third test the ability to take two and timee-part dictation with modulations and more difficuld riythes.

[^9]The Aliforis Freshan Test in Inasic Achierement ${ }^{\text {I }}$ is a recent test which is dooigned to be ased. with inusic majoxg at the college entronce level in neasuring the auditory-visual discrimination of melodic, hamonic, and mindmic elements and idions. The stadent ia to select from masical notation itens which he hears played on the piano.

The elements and idions are derived from the foree basio constitum ents of masic, i.e., melody, harmory, and rhytha the first test is on melodic olements, the second on melodic idtoms, and so on through hamony and aytimn. The eloment is decined as tho snallest tonal organization that has sigmificance to the masician. The taion uses the elememt in a simple setting. ren six tests are as follows:

1. Trelodic oleneats. Four mo-tone intervels are mitten in notation, the first note of each beine the sano. The one with the examiner plays is to bo selected.
2. Welodic idions. In a melodic group of four tones, the first three are consecuitive From four possibilities in tive notation the correct fourtin note played is to be selected.
3. Harmonic elements. Major, minor, and diminishod choras in fom-part hamony are notated on a piano staf. Wach iton presents four chords, all difforont oxeent that the soprano note is the same in each. Wach chord is played twice, and the subject is to decide which notation represente the one played.
4. Tamonic idions. Three three-chord sentences are notated. In each the soprano and bass are the same, but the alto and tenor are different. The evaminer plays one of these sentences twice; the subject is to select the one played.
5. Rhytumic elements, Bach item of tho whtten score contains four different rhyitmic naiterns of one-beat duration. One of these pattems is player tiree times in a scalewise nelodic figuration. The patiem used is to bo selectec.

[^10]6. Rhythoio Idions. Wach ttem conslsts of a tro-beat combination of two different riythino figures. The combination patterns are played three times in a scalewise passatge. The subject Is to select the pattern used from the three notated posaibilities.

Homs are provided for comparisons on mational basisy by region, and by type of school. The reliabilities for partmscores and total score
 total, .88. No informaion is given regarding vallazty. It is apparently based upon an analystis of the material found in freaman theory and harmony textbooks, and upon the opinion of music instructors as to what music students ought to know.

## Tests of 保ical Aptitudes

Teste of musical aptitudes are sometines classified as tests of telent or of masicality. Talent is defined by Schoen as the inborn capacity to develop musical production; ${ }^{1}$ musicality ts the eseential quality necessary for mustcal sensitivity or artistho performance. The two are related in that the greater the degree of musicality the finer will bo the perionnance In muacal value, provided the necessary talent is present.

Tests of talent measure those spocific elemental capacities, such as auditory sensitivity to wariations in pitoh and rhything which are considered to be neceasary for musioal production; tests of masicality are more concerned with the attributes of the musically gifted person.

[^11]Both talont and masicality are necessary for misical success, but since musiculity ts the prine condition for artistic masical perforance, tests of masicality can be nsed as tests of talent. Fron the standyoint of talent mustical aptitude tests ancwer the question: Given the necessary prexequisites for technically artistic rendition, is the person sufficiently endowed by nature with musicianship so that his periomance vill also have a satisfactory degree of musically artistic value?

The assumption is rade witin these tests that they indioate native capacity rather than the resulis of training. Such an assumption is true, according to Schoen, in that "it is tina that the ciegree to walch a person profits br experience along a certain line deponds upon hie natuve endownent." ${ }^{\text {I }}$

The most widely lenown of the musical aptitude tests are the Seashore leasures of fusical Talent. In the revised edition there are six subsests, recorded on phonograph records. There are toro forms, $A$ and B. Series a is for unselected groups and is designed to bo used for genm eral testing purposes. Series 5 is Intended for the more intensjve testIng of musicians and studente. The tests will not be described in detail since tiney have been extonsively used and have beea moch ciscussed in tine Litcrature. Briefly, the subbests include: (1) pitch, (2) Ioudnest, (3) time, (4) timbre, (5) ruyth, and (6) tonal memory:

ITbid.s p. 160.
${ }^{2}$ Carl E. Seashore, Don Jerrie, and Joseph G. Saetveit. Seashore Leasuros of Luefcal Talents. Fducational Departnent, ROA Victor Division, Kadio Corporation of America, Canden, N. J*, 2939.

In construatine hia mowaseg, seachors malysed muacal tulent theoretically into components or basic capacities and selected for study those which conli be measured objectively. Theso are all masures of auditory percoptian and acuity, but the clann is not made that they represent all the wisnifigant componenta of maximal talont.

Seashore raintained that each of the subteste should ba noparatoly valideted against different kinds of specialized musical activity, such as porformance on the pitich tent by players of stainged instrunents; who rust be able to make fine gradations or adjustments in pitch. guch walidation studies have not been mace, homever, reliance boing placed upon intemal consistency as the truest and most scientific evidence of validity. Ay intomal consistency is mant the isolation of the factor being measured and the control of all other factors in the measuremont. The assumption is that persons who score high on the subtests ought to do cqually moll with these "talenter in a musionl situation, other condicions being equal. ${ }^{1}$

Nevertheless, the practical value of the tests as predictive devices requines some overal validation against external oriteria of musical muccess, such as tazoher ratings on musicality and musical achievement. Hang studise of the raithity of the teats using nuch criteria have been natis and reported in the literature. The rosults senerally have indicated that these tests are of doubtiful validity in predicting various levels of masical talent. Their chief value is of a negative

[^12]noture: the tesis neved those pergons whose and tory percoptions are कo deficient that they conld not successfuily participate in tiv formal sbudy or perfomanco of maesic, ${ }^{\text {I }}$.

The reliabllity of the Beashore Keantes has Likowise been subu fected to considerable roseart, particalariy tho 1919 terston. These
 testa of pitch and tonal momy have sufficimb reliability for certain diagnotic purposes, and that the othere must be used with cautione the revised version, with sdult anjocts, hen had wotter rojubilitios roportod. ${ }^{3}$ These, by tho mhitwhen nethot, are given as follows for Series $A$ and $B$, pitch, .88, .78; 1oudness, .68, .77; tine, .75, .70; timbro, .74, .72; rhyth, .62, .72; tonal memony, .88, .89.

Tests similar to the heashore twasures in thein moarurenent of
 several of tios subtosts measure troined aulities, and further difior in that actual musical material is used the whole battory consists of ten tests, each recorded on a phonograph record. Those sirallan to the

Prank 3. Treenan. Theory and Eractice of sycholocical Festing.


2R. R, Farnoworith. "An mistorical, Critical, and Enperinental study of the Seashore-kwalvasser Test Battery " Genatic Psychology Konographes, 9 (1931) 291-393.

3J. C. Saetve:t, D. Leris; and C. B. Seashore. Revision of the
 Ahother revicion of these tests is in preparation, and is to be published by The Frychelogical Corporationg Now York.
 of Dirootions. Garl Fischer, Inc. New Yorl, 1930.

Seashore measures are: (1) pitch discrimination, (2) Intensity discrimination, (3) time discrimination, (4) rhythmic discrinination, (5) quality discrimination, (6) tonal memory.

The remaining tests attempt to measure musical feeling and appreciation, and tonal and riythmio inagery. These are described in more detail:
(7) Tonal movement. Each item consists of a four-tone tonal pattern which requires a fifth tone for completion. The subject is to judge whether this fifth tone should be higher or lower than the fourth tome given.
(8) Melodic Taste. In each item two melodies of two phrases each are played. The first phrase of each melody is the same, but the second phrase is different. The subject is to judge which of the tro melodies has the more appropriate second phrase.
(9) Pitch Imagery. Tonal pattems are given in notation, and the subject decides whether they are the same as or different from those heard on the record.
(10) Rhythm Imagery. The subject is to compare the Visual rhythoic patterns with those sounded on the record.

The reliability of the tests has been found to be amall. Farmsworth ${ }^{1}$ concluded that they were too unreliable for inditidual prediction, with the possible exception of the test of tonal movement, and that they should probably be limited to studies of group differences.

Their validty is questionable. A number of studies comparing the tests with grades in music courses and musical perfomance have shown low; and in some cases even negative correlations. Nursell ${ }^{2}$ has attributed the

[^13]lack of relationship between both the $\mathrm{R}-\mathrm{p}$ and the Seashore tests and criteria of misical ablity to the fact that they deal mataly with seneory capactites, which are relativaly unimportant Ingredionts of musicality. The low comelations mith course marks may be due in some measure to the lack of reliability of the oritoria, however.

Another test of mulcal talent ls the Drake battery. There are four subtesta:
2. 解atcal memory there are twelve ttens, each consisting of a two-measure melody, which is repeated with a change in key, thes, or note, or there nay be no change. The subject is to dotermina by hearing waich type of change has occurred. If any. The melodies are arranged in order of inoreasing afficulty. The test, for which thame are two forms is interded for use in grades 3 to 12. However, percentile noms for vanlous age Ievels from 7 to 24 are given in the instruction menual.

Dykema and Gehreene ${ }^{2}$ rogard this as an achievemant test, but Drake considers it a test of musical talent. The assumption is that musteal memory is an important indication of musicality.
2. Interval discrininetion, Pach item consists of two intervale, the oubject deciding whethor the second ie longer or shorter (larger or smaller) than the firet. The second interval is always played above the firet and o11 intervals are ascendinge. The intervals increase in difficulty during the firat half of the tost, and decrease in the second half.
3. Retentivity. This is a complicated test which places great emphasts on memory. The filrst part of ach item congists of a masioal Interval, netronome beats, and a three-note sequence. The second part consists of several mutical intexvals; more matronome besta, and a single note. The

1n. W. Drake. Wour Nem Tects of Mhsical Talent." Joumal of Applied Psychology, 17 (1933) 136-117?.
${ }^{2}$ Peter ${ }^{2}$ Dyteral and Karl Cenricens. The Teaching and Administration of High school Musio. O. C. Birchard and Co.g Boston, 1951. P. 372.
subject is to judge whether each of the intervele of the second part is shorter or longer than the interval of the finst part, whether the matronome beat is faster or slower than that of the first part, and whether the single note of the second section is the first, gecond, or third note in the sequence of the first part.
4. Intuition. This test attempts to measure feeling for key conter phrase balance ant tine balance. The assumption is that feeling for musical form is an important aspect of musical telent. Each itera consists of two melodic phrases. The subject is to judge whether or not the second phrase is atatisfactory in relation to the inst. Each phrase is supposed to test one of the above three feelinge, or "intuations."

Reliabilities are given by Drake as: .93 and . 85 for musical memory; . 74 and .43 for interval discrimination; and .76 mad. 53 for retentivity. Validities baged on teacherst rankinge of studentio are given as .67 and .54 for musical memory; .58 and .42 for interval discrimination; . 54 and .38 for retentivity; and 36 and .35 for intuition.

The firgt test, Rusical Memory is the most sativenactory of the foury and is available separately. ${ }^{2}$

Another masical aptitude test which emphagizes the inportance in musicality of the ability to diacriminate the quality and tonal relationship of intervals is the test by Madison. ${ }^{3}$ Euch Item of the test consists of a short series of two-tone harmonic intervals, such as found in actual musical material, one of which is different in its sound from the rest.

[^14]This the subjeot is to identify. The dasorininations raye from very eacy to very difficult. High xeliminility hos been reported, and oomparisons with cwerall criterie of musicolity have indicated high rilidity. The sohoen battery of testo ${ }^{2}$ was designed to measure general rucicianship as indioated by aesinetic reaponse to the standtural elements of music. There aro throe tests: those of reletive pitohy towal sequence, and riythm. Senetivivity to these elements is considsred to be evidence of musicality, and the degree of davelopnent through brainine is considered to be dependent upon native endoment or exuipnent.

1. Relative pitch. Paire of intervaze of freveasiric difficulty are to be compared as to size. The subject is to judge whother the second intarval is lerger or analler than the Eirst.
2. Tonal sequence. Thie test conslsts of four womphrase melodies Erom standard compositions. Zath relody has three altermate second plurases and endingsy as well as the original. These are arranged th orcier from most orvious to most complex: and the subject is to rate the relative apsopriateness of each ending phrase. Criteria used in rating the oxcellence of the melodies (impression of comploteness of the tonal sequence) mere conerence ("belonging-togetherness"), balance, untby (and variety), and finality or effect of conjuth to rest.
3. Rhytho Each them convista of a pair of thythmic pluaees, with each phrase containing two rlythaic patterns. The two phrases are allke except for a slight change in the. duration of one tone of either the fierst or aecond patm term of the second phose. The rhytings, chosur from actual music, but presented in monotone, are arranged in order of increasing difficulty as to leagth, ginilarity of the patterns, complexity of meter, and slightness of variabion. The subject is to judge whether the second

[^15]phrase is different from the first, and if different, whether the change occurs in the first or second pattern of that phrase.

On the basis of comparisons between test results and teachers! ratings of ten music pupils on six factors of artistic musical performance, the author claims to have established the validity of the tests. However, the test results included those made on five of the Seashore tests, and therefore the predictive value of the Schoen battery alone has not been clearly established. In a separate study by Drake; ${ }^{1}$ the tonal sequence test, the only one of the Schoen battery which was. Included, was not Found to be satiafactory in both validity and reliability.

A series of tests by Lundin ${ }^{2}$ attempt to measure some kinds of musical ability which had not been investigated previously, and which are Judged to be important behaviors of musicians since they are emphasized in theory courses. There are five teats, the first fom being recorded using a Hamond organ, the fifth using a plano.

1. Interval disorimination, Each item consists of two pairs of melodic intervals. The second interval may be the same size as the first, but the pitches axe transposed. The subject is to judge whether the second interval is the same as or different from the first.
2. Nelodic transposition. Pairs of simple melodies are given, with the repetition of the melody always transposed to a different key. The gubject is to indicate whether the second melody of the pair is the same as or different from the first.
3. Mode discrimination. Pairs of single chords are given, with the second chord transposed to a different key. The subject

[^16]is to indicate whether the mode of the second chord is the same as or different from the first chord.
4. Melodic sequence. Each item consiste of four nommodulating melodic sequences, with the first three following the pattern or varying from it. The subject is to indicate whether the last pattern is the came as or difierent from the first three.
5. Rhythmic sequences. This is similar to the melodio sequences, except that rivthmic patterns ame used. The ryythas appear in a melodic sotting as found in actual nusic.

The test3, according to Lundin, apparontly maasure behavior more typical of musicians than of unaelected individuals, since the difference between the moans of those Eroups, for each test and for total scores, was found to be very signiticent. The tests were validated against ratings of music students in music courses and in perfomance. Total tost scores yielded validity coefficients of . 70 for melodic dictationg . 70 for harmonic dictation, 43 for writton hamoniaktion, 65 for general ability in theory, .51 for performance, and .69 for a spectal category, the sum of ratings. The individual tests are also reported to have high correlam tions in general with the seme exiteria.

The reliability coefficients, computed by the split-helf method for the masician and unselected groups separately, are considered to be high enough to allow the tests to be used for general predictive purposes, With the exception of the test of nove discrimination with unselected subjects. Total scores reliability coefficients were found to be .89 and .85 for the musicians and unselected groups, respeotively.

Other tests in wisch musical discriminations are to be made are those devised by ortmann for students entering the Peabody Conservatory of Music. Like the geanhoye min tests, their usefulnoss appears to be linitu to gtudies of group differences. Thena are seven tests:

1. Pttch discrinination. Pars of tones are given, and the subject is to jucige whether the second tone is higher or lower than the fitrst.
2. Fitch manory. A sertes of tones is given, the flrst of which is to be memorized. The subject is to judge other tones as being the same, hither ow lower than the first.
3. Time discrimintion. Pairs of Piva-cliok patterns are given. The subject is to judge whether the second pattern of each pair is more even or mone uncvea tham the first.
4. Musion. Tonni conbinations Tamiag from two to four notes. are given. The subject is to judge the number of tones in each combinetion.
5. Rhythic momory. Wach ttom of a beries contains a fitemoliok riyythoic pattern, against which other patterne are to be compared as being more syon, fore unaven, or the same.
6. Melodic memoty Palts of melodies ranging from two to six notes are given. The subject is to indicate which tones of the second melody of a pair differ from the tones of the first melody.
7. Hamonic momory. 太eries of chords are giveng each beginning With a standert chord. The subject is to judge whother several of the chords are the same as or different from the bogiming chord.

Retest reliabilitity of total ecores was found by Petran ${ }^{2}$ to be .80 with 500 students. The highest single test rellability was 86 for Pitch

[^17]Memory and the lowest was .30 for Melodic Memory. The validity of the tests was estimated from an analysis of the accumbated data from the tests at the end of a ton-year period, these data including the pupils' success in music. One of the purposes of the tests was to determine differences in test performance between pupils majoring in piano, organ, voice, composition, string instruments, and school music. No striking differences were found except that string instrument students scored higher than the other grouss on pitch discrimination, whereas composition students averaged higher scores than the other groups.
$A$ Test of lusicality ${ }^{7}$ by Gaston has been designed primarily as a musical aptitude test, but parts of it require knowledge of notation and thus measure the degree to which musical capacities and abilities have been developed. Information as to the pupil's musical background and attitudes toward music provide additional data for diagnostic purposes. The test, which is recorded using the piano, is in four parts.

1. Tone discrimination. A single tone is heard, and then a chord. The pupil is to indicate whether or not the single tone is in the chord.
2. Melodic notation. Short melodies are given in notation, and are to be compared with those heard on the record. The pupil is to indicate whother the melody heard is the same as its notation, whether a note, or whether the time has been changed.
3. Tonal movement. Melodies are heard, with the last tone omitted. The pupil is to indicate whether the last tone should go higher or lower.

[^18]4. Welodic memory, Each itom consists of a melody which is heard once and is to be remembered. It is then repeated sevaral times, with a change in note, or in time, or remaining the same.

The splitwhalf reliability of the test was found to be . 88 for grades 4 to 6, "88 for grades 7 to 9, and 90 for grades 10 to 12. The validity of the test rests upon the association between the teacheris ovaluation of the musical personality of the pupil and that pupil's score on the test. The evaluation conslsted of plecing each pupil in one of five categories ranging from "poor" to "excellent" Since this did not result in a numerical score, the association was teated through the analysis of variation by method of ranks, resulting in a chi-square correlation. The ranik correlation ratio was found to be .52 for erades 4 , 5, and 6, and. 67 for grades 7 to 12. Norms are given according to grade (4 to 12) and age (9 to 18) for both sexes, based upon 5,896 test results.

Another recent aptitude test is that by whistier and Thorpe. In this test, as in some other recent tests, musical examples are used, rather than mechanical devices of a nomusical nature which were comon in earlier musical aptitude tests. Ausical examples of pitch, rhythand melody are to be played by the examiner. The test has five parts:

1. Rhythra Recormition. There are ten items, each consisting of two ryythas, each two neasures in length. The student is to recognize whether the second riythm pattern is the same as or different from the first.
2. Pitch Recognition. There are ten iteme. A tone is first played at an established pitch; in the melody of four
$1_{\text {Earvey }}$. Whistler and Lovis P. Thorpe. Kusical Aptitude Test: Sertes A. California Test Bureau, Los Angeles, 1950.
mensures of thirteen quartar notes thich follows, the student is to determine the number of times the established pitch ocours.
3. Nolody Recogntition. Each of the twenty-fite items consists of two nelodies, each two measures in length. The student is to decide whether the second melody is the same as or different from the first.
4. Pitch Discxinination. There are Iffeen items, each consisting of two played chords. The student is to decide whother the second chord is the sans as the itwst, or higher, or lower.
5. Advanced Rhythn Recogntion. The items are similar to those in Pard 1 , but more difficult.

Reported reliamifties are migh. Valiaity was investigated by correlating teachors' estimates of instrumental talent and of vocol talent With part and total scores, and such factors os whether the student at the time of testing had played an instrment at least a yeary or whether be was a member of a band, orchestra, or chorel group. The reported correlations, obtained by the use of the otis unlike signs method, range Irom . 21 (total score with membership in chorus, choirs, or glee club) to .78 (total score with teachers' estimete of vocal talent)when corrected for attenuation. Intercorrelation data for part and total scores are also reportec.

For purposes of scoring and preparing a diagnostic profile, the two sections for pitch are combined, as are the two for rinthn, thus making a total of thee subtest scoresmpitch, rhythm, melody. Tobles of norms in decfles for the trree subtosts and the total test are given for each of grades $4-8$ and for combined grades 9-10.

There se several teste wioch are primarily concermed with musical sensitivity or feeling rather than with special capacities or aptitudes.

Such tests are less objective and may be considered to be exparimontal in natare。

The Gadence and Fhrase Tests by Loweryl are an sxanple of this type of test. In the cadence test pairs of cadences in various positions are given, using a piano, and the subject is to indicate whether the second cadence of a pair is more or Mess final that the first. In the phrase test tonal sequences are given, with a possible difference in phrasine. The subject is to indicate whether the second sequence is the same as or different from the fismet.

In the Test of HeIodio and Hamonic Sensitivity ${ }^{2}$ Kwalwasser attempted to measure the abillty to discriminate good from poor, short melodic prom gressions, and the ability to judge good from poor, threemohord hamonic progressions. Established musical procedures were the oriteria for judging the quality of both melodie and harmonic progressions.

The Gregon Musical Discrimination test attempts to measure the ability to discriminate good fron bad music. Sach tem of the testy which is recorded on recoras, consista of brief paired musical selections, one of the pair being as written by a recognized composer and the other being altered in riythat, harmony or melody. The subject is to indicate which

[^19]of the pair he prefers, and also whether the difference between the two is in the rhythm, hamony on melody. In the first report of the test ${ }^{1}$ each Item constated of the correct version and three antered, but the simpler forse ${ }^{2}$ with only two versions was found to be more satiafactory.

## Sumary

The teents which have been reviewed may be considered to be reprem sentaitive of the principal types of tests of masteal beharior. The oarlier tests of musical aptitude, particularly the Seachore series, have been criticised as being acoustictil wether than masical in character, measuring sensory capacities rather than reactions to the dymic patterms of tone which characterize masical parception. Thay asswied, without any adequate evidence, thet these devices would be ditanostie of a genersal musical capacity. The more recent tests have used more yousicel materials and more musical situationt in which to measure thoge factort considered to be essential to musloal suecess, and are therefore finding move ready accoptanod by misic ocucatore.

In general, attempts to valldate musical aptitude tosts against ine oritemia of success in musio studies and in music as a profession have been disappointing. Thit is more true of the earller tests; the Validity of the more recent tasts as judged hy extermal criteria has inproved.

[^20]The validation of music tests has aldays been a troublesone matter, due to the difficulties inherent in the analysis of the couplex factors involved in masical personality and in musical success. The choice of criteria of success also involve difficulties. For instance, anong the most used oriteria are course maris and teacherst estinates of talent, but these criteria are often Iow in roliability. Course maxks may sometimes even be assigned on the basis of factors that have nothing to do vith music. The question tiso ariess as to whether the tests are measuring factors more closely related to achievement than to aptitude.

Heny of the standardized achievemont tests appear to be quito similar in content. Some are mislabeled, being merely tests of knowledge of the rudinents of masic, and are very atomistic in their measurement of ivolated bite of knowledge or skill.

The teste of achievement in musical skilley except possibly those of sightmsinging ability, appear to be limited in thair application, both as to school level and as to purpose. Teats of achievement on the college level are soarco. The Aliferis test is one of the fey designed for that level, but its usefulness as a diagnostic device has not yet been estabIished. Due to the psychological controls which are applifed, the items. while utilizing recognizable musical elements and iajons, are not lifelike in their application. Although it is true that musicians mast be able to correlate sound with notation and notation with sound, such ekill is nomally not applied to selecting one correct notation fron a number of differently notated possibilities. The tame observation would apply to other tests which utilize this technique, for instance, the Knuth test.

These bests, ani other achicvenent teste whoh have been revtewed moy in teremi fustly be criticized an being too often characterized by test sithations which are milike actuel matcal situational The question arioes as to the importance of whatever knonledge of shill is kuing tested. Wany of the esste reviemed have sercal assimptions in conmen, and athlum severol tecontques in comon whioh had some amacerion in tho present atudg. All the achievement toste ossume a kowlegige of musicel mimble to be important, as well as time and key signatures, and note
 chords.

In the achinvenant bestry as as in sone of the aptitrade tests, the recogntion of familiar melodies trom notation is an ofton used tecinique, and presumably sequires sone use of aural inagery. Aural inagery is further stressed th the technique of whing tonos and tonst patterns from dictstion. From the techaques of wecognang fanltar melodies it
 dies from theit notatlon. Several bestoy such as those by Alizexis and Shuth, deal both with the actual perception ond awor hearing of notetion through the technique of selooting the cotrect notaton from amont varlous alternatives.

Perception and jmagery are also deslt with in mother techmqus, that of detecting surtimat or tomal errors in the sown of notated patm terns or melodes. From the it would be only a stey to the retection of exross in the sound of notated intervals and chords as thoy eacur in the actual hamonic progressions of real masio. It ig probeble thet
informal efforts heve already been made to measure this skill in some manner similar to that undertaken in this study: but they have not been reported in the literature.

## CHAFTER III

THE CONSTEUCTION OH THE TEST

The Rationale and General Procedures
"The way to measure the extent to which an educational objective has been realized 1 s to aim at it directly whenever possible, rather than to infer its existence from the indirect measurement of sometining else. 1 Very often in music tests the extent of a capacity or skill can only be inferred through the measurenent of processes which are assumed to be related to the capacity or skill. In the present study the attempt was made to measure as directly as possible the ability to read and hear the tones and hamonies of music as it functions in a meaningful situation.

Since intervals and chords are functioning elements in a complete context of other tones and chords, a score reading test should provide complete patterns of a phrase or more in which the tones and harmonies can function meaningfully. The detection of discrepancies between the notation and the performance of tones and harmonies of actual masic would thos be a meaningiul activity, and, as used in a working situation such as a simulated rehearsal, would provide an indication of functional scorr reading ability.

[^21]Burns ${ }^{1}$ and Hoggard ${ }^{2}$ have both suggested similar techniques in measuring the ability to read the score. Since Burns found that the obility to hear the score by looking at it was the most important general harmonic technique used by choral leaders, he considered that measurement of the ability to hear the score could utilize the possible techniques of error detection or of hamonic dictation. Harmonic dictation is suggested on the assumption that the ability to mrite what is heard is evidence of the ability to hear what is seen. This technique, horever, was not considered to be feasible in the present study for several reasons: (1) the basic assumption would probably have to be validated nore substantially than it has yet been; (2) it places a premium on the ability to write music, the value of which is questionable; (3) It is a cumbersome and timemconsuming method; (4) its administration would be difficult; and (5) it is not a functional situation for music directors.

Regarding error detection, the following suggestion seems more feasible; "Chords would be played for the listener who would hold before him a copy of the score. The listener would discover and mark chords which were rendered as they appeared in the written score, and those

[^22]which were not. For those that were different he would indicate detaila of difference. ${ }^{1}$

Hogeard has sugested stinlar techuiques of error detection to tost silent reading: The score contains wrong notesp pianist plays correct ones; the score is correctly mitten but the pianist makes mide thakes; ons porson sings, the class listens, reads, and cheoks variations from the score; a quartet sings; the class watches for any discrepancles. ${ }^{12}$

In test construction objectiviby is gained by defining caroftully the behavior to be tested, and then ereating situations in which responses Will be confined within definite and narrow linits. It appeared feasible in this test to linat the type of response stratety to detection of the same type of emors, and in the sumedim lest it affect the response.

Begarding the type of error or discrepancy to be discovered, the only defensible one fror the standpoint of practicality is that one in which the error occurs in the performance and not in the notation. To be sure, an occasional mismint nay appear in a ecore, but by far the greater number of errow occur in persomance. This Jimited the type of exrors which could readily be introduced into the test, but othexwise the test could be criticized justifiably as being nonfunctional or unrealistic, and having more academic than practical value.

In comrection with the type of error to be introduced, the choral medium was chosen, among other reasons, because of the fact that choir

$$
\begin{aligned}
& 1_{\text {Burns, pp }} \text { 102-103. } \\
& 2_{\text {Hoggard, p. }} 24 .
\end{aligned}
$$

singers are often poor sight readers and comit mary exrors of all desm criptions. From auditions of 1000 applicants for the Fred Waring organization, some startilng results were found regarding their masical ability. Table 1 shows that only 1.7 per cent were classed as highly prom ficlent readers and 12 per cont were rated as muslcally illiterate.

Table 2 shows that even the majority of persons with some college and conservatory training in music do not rate very well in reading ability, since almost 11 per cent were rated as having only slight ability to read, and almost 69 per cent made many mistakes and needed help to continue. Thus, in this test, the erross nay readily be justified as being caused by poor saght reading on the part of the singers.

Some errors may be caused by singers who nalsjudge the relative size of intervals, or who cause changes in modallty due either to mism judging intervals or to a weak sense of tonality. There are, of course, those singers who sing "by ear" only, and they can provide all types of errors, such as a thitd in the harnony when it should be a fifth.

However, the test about which this study was concerned was to measure the ability of masicians in the simultaneous reading and hearing of tones ond harmonies, and therefore the errors had to be such that aurel and visual faculties were called into play. Obviously, if some tone or chord is discordant or out of character with the general flow of the music or with the general chord context, the merely musical person. as distinct from the person trained to read music, conid detect the error. Many errors are of this type, but if the muste director relies only upon a generally euphonious sound and flow in the music he very likely will overlook nany more subtle errors.

## TABLE 1

## OLASSIFICATIOL OF 2,000 APRLIGANS ACCORDTHO TO READIMG ABILITY ${ }^{-1}$

| Nating | Description | Per cent |
| :---: | :---: | :---: |
| I. | Excellent, independent readers whose skill paralleled that of professional orchestra musicians. | 1.7 |
| II. | fead simple material well but confused by unusual intervals key changes, and complex riythms. | 11.2 |
| III. | Read fairly well but made many mistakes; helpless mithout the paano. | 45.8 |
| IV. | Demonstrated very blight ability to read | 29.3 |
| V. | Could not read at all. | 12.0 |

[^23]
## mablis 2

 OIN THE COLLEGE OR COMSERVATORY LEVEI

| Raturg | Description | Per cont |
| :---: | :---: | :---: |
| I* | Excellent, independent readers whose oldil paralloled that of profersional orchestra musicians. | 2.6 |
| II. | Read slingle maderial mell, but confuced by unusual intervals, key chenges, and complex rhythens. | 17.7 |
| III. | Read fairily well but made many mistalseat helpless without the plano. | 60.9 |
| IV: | Domonstrated very alight ability to read. | 10.8 |
| $V$ | Could not read at all. | 0.0 |

$1_{\text {Adapted from }}$ Hoggard, Fig. 3, p. 4.

The test was intended to be functional, and therefore the errore had to ba introduced in as natural and as logical a manner as posslble. This considexationg together with the necessity for introducing errors which could be detected only by the condinated function of eye and ear; limited somewhet the sampling of chord vocabnamy with which the theory teechor 1 s concemed. and also the serpling of choral Itterature with Which the music director is concerned. Howevery an attempt was made to Introduce errors into a representatite sampling of, at least, the commonly used chords Fon thia purpese a temulation was kept of tho intervals and chords into mhich the erroxe of periomance were introduced, on the basis of quality or moda (major, minor, oto.) and on the beris of function (tonde, dominant, ete.).

As to the representativeness of the choral literature, an attempt was made to cample that literature which the average high school, college, or churoh choir might uss. gue to the necessity of restricting the length of the test, it was not possible to samplo thoroughly this vast field. Several considerations predominated in the selection of the material: its availability Sor, and litelthood of genoral use, and the prodominasce of definite tonal content, i.e., diatonic music, with which the sverage choir is probably more fand lise, and the predoninance of three and four parts. Since the traditional. study of bamony is based upon the common practice of composers of, roughly, the oighteenth and nineteenth
centuries, ${ }^{2}$ particularly Bach, ${ }^{2}$ and stince the music of Bach and his contemporaries is still extensively used by achool and church choirs, omphasis was placed upon. that type of music:

The test, then, consisted of a number of excerpts selected from various types of readily avallable choral Ifterature such as the average choir might use. Bach excerpt was a phrase or more in length, aufficient to provide pattern and meaning to the tonal alements. Errors were introduced into the intervals and chords such as might for the most part be logically caused by inaccurate sight reading on the part of the singers, and these errors were to be detected and indicated by the musicians taking the test. In order to keep the test mithin reasonable time limits and also provide sufficient items, nost of the excerpts had more than one error. This ras reasonable, also, from a practical viewpoint; bingers do not necessarily limit themselves to one mong tone per passage.

In order further, that the test mifht approxinate as nearly as possible a rehearsal situation, and thereby be of functional value, it was considered advisable to rapeat each passage. This decision was based on the reasoning that in rehearsing a selection the drector is listening to make sure that the choir is perfoning the notes correctly. He thinks he detects an error; but often cannot be quite certain since music is a complex matter and passes quickly. So he may call for a

1 Falter Piston. Harmony. W. W. Norton and Co.s Inc., 1948. P. 2.
$2_{\text {As recent and as nidely used a harmony textbook as The contra- }}$ puntal Hamonic Technique of the 18th Century, by Allen Irving Mailose (F.S. Crof'ts Co., Neff York, 1947), is based largely upon the music of J. S. Bach.
repetition of the pessage to confinm hie judgnent regarding the errors.
The construation of an achievernent test usually follows certain major steps, as follows: (1) Planning the test, (2) Hriting the test excrcises, (3) Trytng out the test in preliminary form and assemoling the finishoa test after tryout, (4) Detemoining the procedures and prom paring the mamals for adrainstering and scoring the test, and (5) Reproducing the test and sccessory materisls. ${ }^{2}$

Plaming of the test and general procedures in terms of the objectives to bs attained have been described above. In each of the remaining stape there were a nuber of problems to consider.

## Exploratory Form of the Test

In ordex to test the rationale of the manner of introducing errors, and to determine procedures for marking and scoring the test Items, an exploratory form of the test was constructed. since it was anticipated that many of the test items mould later have to be eliminated, a large nuber of passoges from choral literature was selected, and many errors were introduced into the perfomance of these excerpts.

The logio of the manner of introducing errors has already been explainod. In the test form included in Appendix $B$ the errors in perfomance have been indicated by red gymbols. These marks; of course, were not indicated on the regular test forms. Every effort was made to

1T.F. Lindquist. MPreliminary Considerations in Objective Test Construction." Educational Keasurement. E.F. Lindquist, editor. Anericen Council on Education, Fashingtong B.C.a 1951. P. 119.
introduce such exroms a could eabily have been mada by inaccarate gingers. ${ }^{1}$ In aditition, it mas considered essential that these exrors, If fhey were to test piluultaneous readig and hearing, should not be Liable to detection merely wpon the basis of obvious inapropriateness to the musical contert.

## Reproduction of the Test

Bcore: - The extempts includid in the explosatory form tere reprounced by manuseript on mpittot etencils. Care wes taken that the manuseript resemblad actual printod ratation as far as possible, and that the notation of the symbols was clearly legible.

Recording: - - In ordex to insure uniformity of administration, the test wes recondeu on magretwe bapeg using a Magnecordar machino. Since this first version of the west was to provide only a crude estimate of the difficulty of the itensy it was mecorded in a piano medim. Each excerpt was played twice, the repetition befing pleyed as nearly like the ferst time as possible. The time spacing of the passages was gs follows: anproxinately three seconde separated the repetition from tho first performance of each excerpty an interval of approxinately five seconds intervened between the second parformance of the excerpt and the beginming of the next passace.

[^24]
## Method of Marking Errors

This test by its nature seemed to demand that rapid marks be made directly on the score, fust as a person rehearsing a musical group might mark quickly on his score those places which dissatisfied him. Therefore it was considered nost feasible to place a small blank box above every note or chord. A check mark in one of those would indicate a discrepancy between the notation of a symbol and its performance. This could be done quite quickly with little interference in keoping one's place as the musical performance moved along However, it might cause confusion to the eye to see small squares scattered about over single tones within the staves, so all boxes were placed in a stratght line over the staves.

This arrangement did not make provision for the indication of the specific notes within chords which were in error. Since it was not known how easy or difficult the test might be, a further aegree of differentiation betreen musicians in the ability to detect orrors was deemed desirable. Therefore, the directions to the subjects were, on the first hearing of the excerpt, to place a check mark in the box over the note or chord which was performed incorrectly, and on the repetition of the excerpt to circle the head of the exact note which was in exror. With single notes this was unnecessary, but otherwise this provided an exact check on the accuracy of tonal discrimination in the case of hamonic intervals or chords.

Method of Scoring

The method of bcoring was to allow two points for each item (erzor in performance) if the box was checked and the correct note also wes circled, and one point for each then if the box only was chocked, or If that box mas chectred mit the wrong note circled. If the note in emrox within a chord mas circled but the box above that chord mas not checked, two pointe were still allowed, since the finer degree of tonal discrimination had been exercised. In the case of single tones, one point wes allowed whether the box was checked and/or the note was circled, since only one degree of discrimination thas required.

## Correction for Guessing Ceneral Considerations

In scoring the tests the problem arose as to whether the score should be the totel number of items answered correctly or whether there should be a correction for guessing and chance success.

Where the score is only the number risht, the general effect of guessing is to raise the score of the guessers. Test makers have been most concerned with the exfect of guessing on tests involving two possible responses, since if a porson knows nothing about a subject he should still be able by chance to ansmer correctily half of the items of such a test. As knowledge of the subject increases, the less guessing Will be nacescary, but in theory, on the average, half of the items which are guessed will be right, and half will be wronge likewise, on a three-response test there is one chance to be right and two to be wrong. On a four-response test there is one chance to be right and three to be
wrong; and so on.
The generalized fomma for correcting for guessing in multiplem response tests is given ${ }^{2}$ as

$$
S=R-\frac{K H}{n-K}
$$

where $s$ bcore,
$R=$ the number of right responses;
W at the number of wrong responses,
n the number of suggested responses for a single item,
$\mathbb{K}$ the number of responses to be selected or marked for each item.

In this test, many of the musical excerpts contained a large num ber of suggested responses, only one passage containing as few as four possible responses, and five containing as few as five. Nost of the excerpts contained not only a great nuber of possible responses, as represented by all the notes and chords, but they varied in the number they contained. Furthemore, the correct item responses (those chords or notes which were performed incorrectly) within the excerpts varied in number, some passages having only one, and others having as many as ten. This meant, considering only the possible responses to be checked in the boxes, that the formula would have to be calculated anew for almost every excerpt. This would have presented a formidable tasls in scoring. For instance, a musical passage might contain ten boxes over as many notes or chords, these representing ten chances for error in performance, and thus, ten possible responses. Yith one exror in performance (one correct

[^25]response) and two incorrect responses checked, the following substitutions would be made in the above formula;
$$
s=1-\frac{1(2)}{10-1}=1-\frac{2}{9}=.78
$$

In an excerpt containing three correct responses (errors in performance) among twenty-five possible responses, and having two correct and five incorrect responses checked, the formula would read:

$$
s-2-\frac{3(5)}{25-3}=2-\frac{25}{22}=1.32
$$

Obviously, such a system of seoring each excerpt separately and dealing with fractions of varying amount in computing the total score for the test would have been highly impractical. Such a system would be further complicated by applying a correction factor to the weighted score obtained by circling the note in error within a chord. For instance, in an excerpt containing ten four-note chords; there are forty possible responses. If there is only one correct response possible, and one wrong response is made, the correction factor becomes 1/39. Thus each excerpt could have two correction factors applied, one for the wrong boxes checked, and one for the mrong notes circled.

The matter was resolved by realizing that the greater the number of choices per test item, the less important it is to correct for guessing. Usually those test authorities who advocate the use of a correction formula do so only with items having only up to five choices. As already pointed out, this test had onjy a few items (excerpts) having as few as five responses. It is interesting to note that the reliability of corm rected scores has been reported to be not significantly different from
that of uncorrected scores, and that the correction increases the validiffy of the scores only slightly. ${ }^{I}$ Rankings of corrected and uncorrected scores mill be approximately the same, correlations between them being reported to be usually . 98 or .99. ${ }^{2}$

It was therefore tentatively decided to apply no correction for guessing to the test items, with one exception, since there were a large number of possible responses in most of the test items, and since corm rection would do little to increase the reliability or validity of the test. ${ }^{3}$ The exception was the case in which every possible response which possibly could occur had been checked. In this case the correction would result in a score of zero for the item.

Some musicians might consider that it would be as serious an error of judgment or knowledge to make a wrong response on a test of this type, $1.0 .$, call a chord or note wrong when it was actually parformed correctly, as not to identify those performed incorrectly, i.e., the correct response. Therefore, the logic of the correction formula may be examined in more detail.

Scoring by this formula involves the assumption that every wrong response is the result of a guess, that 211 wrong responses are equally attractive or equally likely to be selected, and that therefore the lew of chance applies to the situation. 4
${ }^{1}$ Cf. Trawler, p. 366.
${ }^{2}$ Ibid., p. 366.
${ }^{3}$ If a comparison of obtained scores with perfect scores were dem sirable, then a correction factor would be needed, but this was not beIieved to be necessary.

$$
4 \text { Trawler; p. } 365 .
$$

But there is a difference bebween parely chance gressing and guessing besed on "hunches" or "halritnomledge." The corraction for guassing is appropriate for purely chance guessing, but "no correction onn bs altogether appopriate for guessing based on some knowledge of the iter, ${ }^{2}$ Traxler points out that if a person "oan make intelligent guesses in excess of the correction for guessing, he is receiving only his just dueg." ${ }^{2}$ He also states that. "in generaly most wrong responses are probably due to inadequate lanorledge or abilitys, zather than to mism information or wrong learning. ${ }^{3}$ In a good tect ezoh incorroct response is mide as plausible as possible, so that the subject who does not posi eese the necessary skill or information will. select a wong response.

In this test the exrors in performance to be checked did not sound any worse than the romeining context of notes and chords. But who can say whether one chord sounds as satisfactory os the next one to a particular individual? In musical passage in which all the harmonies sound well there is likely to be a spocial consideration of those herHonles which for some reason do not seem as appropriate as they might. Such special consideration may lead those persons who lack the ability to comrelate the performance with the potation to select incorrect test responses, Such considered responses, which are inherent in a test of this type, are not chance, but are based upon some type of musical consideration, and therefore the formula would seriously overcorrect.

> I Ibid. p. 348.
> ${ }^{2}$ Tbia. : p. 349.
> 3 Tbid. p. 366.

To sumprize, then, it appeared advisable not to correot for guessiny except to penalize with a scere of zero those excerpis, with thair attendant Item or itors, in which every possiole reoponse was checked. However, the wrong responses within each musical excerpt were tabulated for two purposes; (1) to determine whether they formed a pattem which would indicate sonething anss with the constriction or perm formance of the excerpty and (2) to deterinine whether, indeed, there was a close relation between the obtained bcores and those correated for chance. To simplity the corpatation, agrose triel correction of ow was applied to eech urong responee. Thls correction factor of -.5 was selected beceuse it not only stmplifesed the computational procedure, but because it greatly overeatineted the average amount of correction noeded. If, with such a fector, the comelation betreen scores uncorrected and corrected for guessing was still high, then no correction rould need to be appiled in the scoring of subsequent test forme.

In the inctructions to the subjectss, they were adrised that it woutd be an error on their part to mark a box or circte a note which had been performed comrectly. Althoush tast speciallats on the whole appear to favor congidered guessing, it was deened advisable that the subjects be dincouraged fron axtensive guessing.

The Sample and Resulta

The semple for the exploratory test included ton musicians of varying degrees of musicn training and experience, ranging from college freshmen to experienced musiciana. The ample was drawn from one four-
year liberal arts college and one state unfversity.
Table 3 gives the obtained scores and the scores corrected for chance success of each indtvidual. Subjects 1 and 8 were experienced musicians who were known to be outstanding in their work. Subject 2 was considered to be an outstanding music student. The much higher scores which they obtained appeared to verify the hypothesis that better musicians would possess in greater measure that skill measured by the test.

A gross correction factor of -.5 for each incorrect response was applied in determining the effect of guessing and chance success upon the rank order differences of obtained and corrected scores. With small samples the Spearman rho method of computing the coefficient of correlam tion is practical, and where it is necessary to know only the rank orders of the differences rather than the exact magnitudes. ${ }^{1}$

Applying the formula ${ }^{2}$

$$
\rho-1-\frac{6 \Sigma D^{2}}{N\left(N^{2}-1\right)}
$$

the rank-difference correlation was found to be .85. This would probably have been higher had the corrected scores been calculated with the use of the formula given on page 56, rather than with the excessively large correction of -.5 for each wrong response. Therefore, obtaining a coefficient of .85 mith a very small sample was interpreted to substantiate the theory that scoring with or without correction for guessing would rank the subjects in essentially the sane order. This finding, however,

```
Ic土. Ross, p. 238-39.
2Tbid.; p. 238.
```


## TABLE 3

RANKS OF UNCORRECTED SCORES AND SCORES CORRECTED FOR GUESSTMG OR CHANCE SUCOESS, USIHG A GROSS CORRDSTION OF -. 5 FOR EACHI IMCORREGT RESPONSE: EXPLORATORI TEST

| Subject | Uncorrected <br> score | Corrected <br> score | $\mathrm{R}_{1}$ | $\mathrm{R}_{2}$ |
| :---: | :---: | :---: | :---: | :---: |
| 1 | 161 | 156.5 | 1 |  |
| 2 | 126 | 120.5 | 3 | 1 |
| 3 | 71 | 59.5 | 5.5 | 3 |
| 4 | 85 | 68.5 | 4 | 6 |
| 5 | 71 | 61.0 | 5.5 | 4 |
| 6 | 56 | 31.5 | 8 | 5 |
| 7 | 39 | 26.0 | 10 | 9 |
| 8 | 135 | 127.0 | 2 | 10 |
| 9 | 66 | 57.5 | 9 | 2 |
| 10 |  |  | 7 | 8 |
|  |  |  |  |  |

was investigated further with subsequent forms of the test.
Since the exploratory test was intended to give only a crude indication of the difficulty of the test items, and because the performance of the musical passages was in the plano medium, whereas the test in its finished form was to be in the choral nedium, the results of the analysis of the itoms which was made are not presented. In the prolininary form of the test those items from this form were excluded thich were missed by all subjects md appeared to be too difficalt, and those excerpts which revealed any definite pattern of messing were aiso elininated.

Tha data on the representativeness of the charal Itterature included in the test, and of the chordal vocabulary represented in the test items, are likemise not presented for the exploratory form. These data are reserved for the finighed (tryout) form of the test.

## The Prelininary Test: Poxt 1

The preliminary form of the test was in two parts. An item analysis of the test as first conceived Indicated that some easier items were needed, and therefore the second part was prepared. Both parts are included in the appendices. The excerpts and attendant items (errors) are numbered for ease in reforring to the subsequent tables of item indices, and the errors in performance have been indicated in red symbols in the musical score.

## Reproduction of the Test Form

Each musical excerpt was carefully reproduced in legible manuscript on "Ditto" stencil. That this effort was successful was ovidenced by the fact that no subject complained of being mable to read the notes easily.

## Recording of the Test

The test was recorded on magnetic tape, using a Magnecordar machine. The recording was made in a partially soundproofed room which was the best available for use. The results were judged by three competent musicians to be satisfactory for the purpose of the study.

The five singers who cooperated in the recording were voice majors in the musio department of a state university, and recommended by the choral director of the miversity as being among the best available singers. The cost of engaging professional singers for lengthy recording sessions would have been prohibitive.

No attemipt was made to secure singers whose volces would blend perfeotly. It was believed that voices which did not blend perfectly or were not periectily balanced would make the individual paris in the harmony more clearly evident, much in the maner of a mixed moodwind quartet, for exaraple.

In order to achieve clarity of the parts in performance only one voice was assilened to a part, oxcept in unison passages. Besides, the sarie mistake in performance would rarely be made by all the singers on a part at the same time, unless it were a deceptive unison passage which
could lend jumelf to sin nemy ervor.
The simgera mere provided with scoros of tha eacempts reminten to contain the Intended eriors, so that they were to atig only what those bcores indicated, Bach gxcart mas rohearsed tintil the parts were correct and the general effect was judged to be satisfactory for the purpose intencied: Then it was recordea as many tines as deent necessary in order to obtein an adequate reconded version. Host of the axcerpts were recorded at leaet three times, and tom were recorded seven or eight times.

From the many versions of each excerpt the best one mas selgoted after oaverul consideration, with speoial reference to the intended errore, but also with due regard for the best performance of the excerpt as a whole.

The excerpts which were finally selected were "dubbed" onto a master tape by using two machines, ad allowing one machine to play from the speaker output into the recorsing input of the otber. Each excerpt selected for use on the master tape was placed into position to play on the first nachins. Then both mechines wexe started simaltaneously and alloved to run until the excerpt was fintshed.

Tining and repetition of excerpts: - - Opimion obtained from the persons taking the exploratory test andicated that more time was desirm able between excerpts in which to becoine oriented. The time spacing and manner of accomplishment were as follows. Sach oxcerpt was first announced by number, which was spoken through a microphone and recorded directly on the master tape. The master tape was then allowed to mun
for approximately ten seconds and then was stopped. The excerpt was then "dubbed" in, following which the master tape was allowed to run for approxinately five seconds. The same excerpt ras repositioned on the first machine and then dubbed in again on the master tape, thus securing an exact repetition of it. The master tape mas then allowed to run for approxinetely five seconds, and then stopped. This completed the cycle for one excerpt. Wach excerpt vas in turn announced and dubbed onto the master tape in this manner.

Some loss in fidelity may have occurred in transferring the music from one tape to another, but it was not apparent. Three musjcians who Judged the recordings could detect no appreciable diference between samples of the original tape recordings and thoge on the master tape.

The completed master tape was listened to by these same three capable musicians, with particular reference to the clarity of the intended errors. For this purpose, they used scores containing the discrepancies in perfomance marked in red symbols, as in the copy included in the appendix. There was agrement that the intended errors in the performance were clearly evident as deviations from the notation in the score.

Addition of Starting Pitches: - - After the master tape was made, It was decided that it would be desirable and helpful to have the piteh of the beginning note or chord sounded for each excerpt. Therefore, when the test was adninistered, the pitch was givon on a chromatic pitch pipe following the announcement of the number of each excerpt. In the case of chords, the pitch of the lowest note was sounded. The pitch was also sounded again before the repetition of the excerpt.

## Wethod of Moricing Errors

The directions for marking errors were the same as in the exploratory test. On the first hearing of each excerpt, the subject was to place a check mank in the box ovor the note or chord which was porformod incorrectly; on the repetition of the exeorpt the had of the sxact symbol in error was to be circled.

## Methods of Seoring

In order to detormine if there was any difference in the results of the possible methods of scoring, inoludine the effect of guessing, and in ordor thereby to select for further use that method which appoared to offer reliable results in the simplest manner, the tests were scored in three ways; (1) Weighted: scored as in the exploratory test, one point for each box correctily checked, and one point for each note correctly circled; (2) Unreighted; one point only for each box correctly checked, disregarding the circled notes; (3) Correction for gressing or chance success: Fethod 1 with a gross correction factor of -.5 for each box inoorrectily checked, or for each note wrongly circled if the corresponding box had not been checked. Comparisons were made of the rank differences between methods 1 and 2 , and between methods $I$ and 3.

## Questionalire

A preliminary form of the questiomaire was formulated. This was for the purpose of obtaining data from the examinees regarding the folLowing factors: age, sex, school level, amount of music theory
instruction, number of years of piano, voice, and instrumental study and experience, and directing experience. No atternt was made to utilize the obtained data in making comparisons based on the various factors, but the questlonnaire as a whole was analyzed subjectively in order to prepare the subsequent final form in a concise, yet comprehensive manner.

## The Sample

The test was given Individually and in groups to thirtymsix musicians. They were selected as being 2 fairly representative sampling Of the music majors on the undergraduate and graduate levela for will ch the final form of the test vas intended. The group ranged from freshmen to thind-year graduate studentsy with a slight proponderance of graduates over undergraduates. The group also included several highly regarded misicians who cooperated in the study. The students were all summersession studente at a state miversity.

## Comparison of Scoring Methods

The rank orders of the scores obtained by method 1 (weighted scores) and method 2 (unweighted scorss) are given in Table 4. The rankdifference coefficlent of correlation was found to be 997 , closely approaching perfect correiation. This fact, together with the consider ation that scoring the circled notes in addition to the cheoked boxes was laborious and subject to error, led to the conclusion that all subsequent forms of the test would be merked and scored as one point for each

## TABLE 4

RANIS OF SCORSS OBTAINED BY METHOD 1 (HETGHTED) $A N D$ BY METHOD 2 (UNEIGHED): PART I OF PRELTMTMARY TEST

| Scores |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: |
| Subject | Method I | Bethoa 2 | $M_{1}$ | $\mathrm{R}_{2}$ |
| 1 | 160 | 85 | 5 | 6 |
| 2 | 212 | 106 | 3 | 3 |
| 3 | 55 | 33 | 24 | 23 |
| 4 | 159 | 80 | 6 | 7 |
| 5 | 126 | 66 | 11 | 11.5 |
| 6 | 54 | 30 | 25 | 26 |
| 7 | 50 | 29 | 26.5 | 27 |
| 8 | 59 | 32 | 23 | 24 |
| 9 | 75 | 38 | 20 | 21 |
| 10 | 45 | 27 | 29.5 | 29 |
| 11 | 17 | 12 | 36 | 36 |
| 12 | 29 | 21 | 34 | 34 |
| 13 | 39 | 23 | 32 | 32 |
| 14 | 218 | 108 | 2 | 2 |
| 15 | 92 | 51 | 15 | 15 |
| 16 | 139 | 75 | 8 | 8 |
| 17 | 80 | 45 | 17.5 | 18 |
| 18 | 45 | 30 | 29.5 | 30 |
| 19 | 80 | 46 | 17.5 | 17 |
| 20 | 120 | 66 | 12 | 11.5 |
| 21 | 249 | 125 | 1 | 1 |
| 22 | 102 | 52 | 14 | 14 |
| 23 | 117 | 62 | 13 | 13 |
| 24 | 78 | 40 | 19 | 20 |
| 25 | 41 | 23 | 31 | 32 |

## TABLEF 4 (continued)

| Subject | Scores |  | ${ }^{\text {a }}$ | $\mathrm{R}_{2}$ |
| :---: | :---: | :---: | :---: | :---: |
|  | Method 1 | Method 2 |  |  |
| 26 | 46 | 28 | 28 | 28 |
| 27 | 81 | 47 | 16 | 16 |
| 28 | 70 | 41 | 21 | 19 |
| 29 | 50 | 32 | 26.5 | 25 |
| 30 | 20 | 14 | 35 | 35 |
| 31 | 32 | 23 | 33 | 32 |
| 32 | 61 | 37 | 22 | 22 |
| 33 | 133 | 68 | 10 | 10 |
| 34 | 137 | 70 | 9 | 9 |
| 35 | 179 | 105 | 4 | 4 |
| 36 | 153 | 100 | 7 | 5 |

box correctiy ohecired. One point per item still made possiole the deter mination of a wide range in the skill and lmowiedge emanded of the sub jects by the testg apparently the finer disarimination required to carcle the exact notes wad not needed for the purpose of the test.

The rawis ondere of the scores obteined by method 1 and scores corrected for chance success are given in pable 5. To almplify the computation of the correction factory a gross corraction of -.5 was applied to each incorrect response in computing the total score. This mas, of course, grectly overestimating the anount of proper correction. In viem of this, the rank-defference coefficient of correlation of .983 whith was obtained was intexpreted as ample eviance thet it would not be necessary to correct for guessing in subsequent forms of the test, except In the case where all possible responses within an excexpt were checked.

Thes dectsion also buplitied the problem posed by tho jecorded pariomance of the test excerpts, in using non-professional equipment ad non-profestional stigers, and with lisited rehearsal and recording tine; dertain inadequactes in the performance occasionaly appeared. These were principally, occasional, glight deviations from the desired intonation, and indistinct tones due to inporfect balence or other causes. It should bo secalled thet the recordod excerptis mere judged critically ty trace capable musiolaus, but with princapal attention to the intended elrors. There was agreenent by the juxy that all the intended axors Were piatnly evident aumally as deviabions from the correaponding notafion. Stnce there was hencefortin to be no correction for guassing, the remaining context of tones and harmonies could be disregarded for scoring

## TABETE 5

RAMKS OF UNCORREGTED SCORES (OBTATNED BY YETHOD 1) AMD SCORES CORRECTED FOR GUESSTNG OA OHANCE SUCCESS, USING A GROSS CORRECIOM OF -. 5 TOR GACH TMCORRET RESPONSE:

PARTI 1 OF FRHLIMINARY TEST

| Scores |  |  |  |
| :---: | :---: | :---: | :---: |
| Uncorrected | Corrected | $\mathrm{R}_{1}$ | $\mathrm{R}_{2}$ |
| 160 | 14.5 .5 | 5 | 5 |
| 212 | 205 | 3 | 3 |
| 55 | 38 | 24 | 24 |
| 159 | 151 | 6 | 4 |
| 126 | 117 | 11 | 10 |
| 54 | 35.5 | 25 | 26 |
| 50 | 27.5 | 26.5 | 30 |
| 59 | 41.5 | 23 | 23 |
| 75 | 74 | 20 | 16 |
| 45 | 34 | 29.5 | 28 |
| 17 | 6.5 | 36 | 35 |
| 29 | 18.5 | 34 | 33 |
| 39 | 21. | 32 | 31.5 |
| 218 | 206 | 2 | 2 |
| 92 | 78 | 15 | 15 |
| 139 | 122.5 | 8 | 8 |
| 80 | 57.5 | 17.5 | 20 |
| 45 | 21 | 29.5 | 31.5 |
| 80 | 63.5 | 17.5 | 19 |
| 120 | 105.5 | 12 | 12 |
| 249 | 246 | 1 | 1 |
| 102 | 94 | 11. | 14 |
| 117 | 101.5 | 13 | 13 |
| 78 | 69.5 | 19 | 17 |
| 42 | 37.5 | 31 | 25 |

## TABLS 5 (continued)

| Scores |  |  |  |
| :---: | :---: | :---: | :---: |
| Uncorrected | Corrected | R | $\mathrm{R}_{2}$ |
| 46 | 29.5 | 28 | 29 |
| 81 | 67.5 | 16 | 18 |
| 70 | 49. | 21 | 21 |
| 50 | 34.5 | 26.5 | 27 |
| 20 | 9.5 | 35 | 34 |
| 32 | -8.5 | 33 | 36 |
| 61 | 45.5 | 22 | 22 |
| 133 | 119 | 10 | 9 |
| 137 | 126.5 | 9 | 7 |
| 179 | 742.5 | 4 | 6 |
| 153 | 113.5 | 7 | 11 |

purposes. This would toke care of the cicessive guessing that might take place in indistinct passages or those of slightiy questionable intonam tiong since it would be oniy satural that a person with a aypercritical ear might find Iault with the siightest Imperfection and call it an error on whatever grounds. In this connection, it is perhaps justifiable to mention that the principal test with which most musicians are faililar is the Seashore test, th one subtest of which very fine pitch discriminations are to be made. Persons faniliar with this test might assume that similar very fine pitch discriminations were to be made in this test, despite the directions to the contrary, and the result would be an excessive checking of mrong responses.

Reliability

In the process of test construction, the reliability of the preIininary form of the test is not always computed, but it is desirable to do so if the finished test is to atisfy certain standards of reliability.

The most frequently used procedures for determining test reliam bility are the equivalent-forms method, the retest method, and the splithati method.

Since there was but one opportunity to administer the test to the subjecteg, the retest method could not be used. For the same reason, the administration of equivalent forms would not have been possible, even if equivalent forms could have been developed. This was not considered to be feasible, for a number of reasons. Each test item had a different context of hamonies. These different contexts; as well as differences
in key tonality, tempo, rhythm, mood, style, dispositions of the notes within items (chords), and varying time values of the items all contrib uted complexities to determining two exactiyg or even approximately equal halves. Furthermore, the musical excerpts contained varying numbers of items, so that it would have been difficult even to matoh excerpts containing the same number of items.

With the split-half method, the single test is subdivided into two presumably equivalent groups of itens, and the resulting two scores are correlated. Usualiy the tho groups are equated for content and difficulty, but within these limits chance should detemine which items go In which half of the test. Among the procedures used in splitting a test to yield tho half-scores are those of alternate items alternate groups of items, and the first versus the second half of the test.

The nethod of alternate groups of items at first appeared to be the most satisfactory mothod, since the groups could consist of the items in context within each excerpt. However, the excerpts contained an mequal number of items in many instances, and it was not possible to arrange two halves of alternate excerpts having the same total number of items.

Splitting the test in the middle was also not possible because of the unequal disposition of items within the excerpts. Even if this had been possible, the assumption that the two halves were of equal diffim culty could hardly have been met. Furtherg other factors such as fatigue or practice could cause the two parts to differ considerably.

It appeared most practical, then, to use the method of alternate items. This had the effect of making a chance distribution of all the
abovementioned contextual factors. It was, in fact, a most rigorous test for it divided by chance the items regardless of their difficulty. Adkins states that the split-malf method makes the assumption that all the items are measures of a single factory in this case the porer to detect deviations in performance from the notation. If more than one factor is involved, this method gives underestimates of the reliability coefficient. "I It was reasoned that if a ht eg coefficient of correlation Were obtained, that the rationale of constructing the test vas funded mentally sound and that one predominant skill or function was being tested, i.e., that the wast was homogeneous in nature.

The coefficient of correlation most commonly computed is Pearson's product-moment coefficient. ${ }^{2}$ The formula applied to a test when has ben split into halves provides an estimate of the reliability for the half test. The reliability of the full -length test, if it is a homo geneoug test; my be estimated by the Spearman-Brown formula. ${ }^{3}$

The scores obtained for the split halves, using the simplified method of scoring one point for each correct response which it had been decided would be used henceforth, are given in Table 59; Appendix $C$. Using the formula for computing 1 from original measurements, 4

$$
r_{x y}=\frac{3 \Sigma X X-(\Sigma X)(\Sigma X)}{\sqrt{\left[N \Sigma X^{2}-(\Sigma X)^{2}\right]\left[N \Sigma Y^{2}-(\Sigma X)^{2}\right]}}
$$

[^26]the coofficient of correlation of the test halves was found to be .975 . The reliability of the full-Iength test, by use of the Spearman-Brown formula, ${ }^{1}$
$$
F_{\mathrm{tt}}=\frac{2 \mathrm{r}_{\mathrm{hh}}}{1+\mathrm{r}_{\mathrm{hh}}}
$$
where $r_{\text {mh }}$ stands for the self-borrelation of the half-test, was found to be 987 . This result indicated that this form of the test yielded highly consistent results, and was apparently homogeneous in content.

Evidence of Validity

Subjects 1, 2, 34, 35, and 36 (see Table 4) were regarded as musioians of superior musion skills and attainments. The fact that they all ranked high on the test appearec to substantiate the hypothesis that better musicians would possess in greater measure the skill measured by the test.

Item Analysis

The quality and merit of a test depend upon the individual iters of which it is composed. It is therefore necessary, in best practice, to analyze each item . . In order to retain only those that suit the purpose and rationale of the device being constructed. ${ }^{2}$

Test items are customarily analyzed for their level of diffioulty and for the degree to which each item differentiates between the high and
${ }^{1}$ Ibid., p. 492.
${ }^{2}$ Frank S. Freeman. Theory and Practice of Psychological Testing. Henry Holt and Company, Inc. New York. 1950. P. 24.

Iow scoring groups. The latter is referred to as the discrininating porar of the items.

## Iten Difficulty

The difficulty of an item is determined by the proportion of the group who get the number right. Obviously, the greater the percentage of persons passing the item, the easier it is. Gomputation of item difficulty, therefore, involves counting, for each itom, the number of persons who get the item right. This number is then expressed in terms of a per cent.

Several disadvantages to this procedure may occur in tests of the multiple response type where guessing has not been corrected for, and where itens have been onitted or not reached. In this test no correction was belleved to be necessary due primanly to the great number of posaible responses. And since each person had to proceed to the consideram. tion of each musical passage mith ita possible responses as the tape recording reproduced then, presumably each person took the opportunity to consider every item. Therefore there were no items omitted or not reached.

The difficulty of a test (and hence of its component items) has an important bearing on its value tor its purpose. A test or iten that is below the level of abilities of the poorest aubjects can be of no value at all in iiscriminating anong the subjects. Similarly,

[^27]a test that is too disficult even for the ablest subjects can give no help at all in predicting which eubjects are superior. An item that is passed by 99 and failed by 1 out of 100 subjects makes only 99 (99 x 1) disoriminations, and one passed by 98 and fajled by 2 makes only $196(98 \times 2)$ discriminations. On the other hand, an titem passed by 49 and failed by 51 or an item failed by 5l and passed by 49 makes 2,499 ( $51 \times 49$ ) discriminations. An 1tem passed by 50 and failed by 50 makes $2,500(50 \times 50)$ discriminations. The single iten that best indicates the level of ability of an individual. is one for which his probability of passing is. 50. . . . As a general rule, the average difficulty of the items in a test shoula correspond to the average ability of the subjects; i.e., the items should be such that, on the average, about half of the subjeots will answer them correctily. If, horever, the test is to be used to belect only a few outstanding individuale, it should be mach more difficult; and if it is to screen out only a ferr extremely poor subjects, it should be moch easier. Nore specifically, if it is desired to select out the top 30 per cent of a groups then the difficulties of the fitems should ciluster around 30 per cent. If, homevery it is desired to plsce the whole group or a sizeable portion of it in rank order with reference to the abilities boing tested, the difficulties of the $i t e m s$ should be spread over most of the range within thich it is desired to discriminate and ordinarily should tond to oluster at the 50 per cent level for the group in guestion. ${ }^{3}$.

In this test it was desired to neasura every individual in the group with as much accuracy as possible, and therefore items from 211 Levela of dfficulty were sought.

The results of the analysis of the difficulty of each individual item are given in Table 60, Appendit 0 , in the colum headed Her cent passing, total group. ${ }^{n}$ 做th thirty-six subjects, the possible levels of affficulty vould be represented by proportions of the total group getting an Item right ranging from 36:0 to 36:36. No items Fere found to have been missed by all subjects, and likemise proportions above 80.5 per cent
(the ratio $36: 29$ ) were not found to be represented. A clearer
$1_{\text {Adktisis }}$ p. III.
understanding of the relative difficulty of the test items nay be had by examining Table 6, in which the number of itoms found at the various levels of difficulty are Elven.

Examination of the item difficulty led to the conclusion that the items were spread over many of the possible levels of difficulty in a fairly adequate menner to measure either individuals or groups. However, there was a preponderance of itens in the lower per cent groups, indicating that the test was on the whole rather difficult.

Item Discriminating Pover

Although the difficulty of an item is a very important factor in the effectiveness of a test, mere appropriateness of difficulty is no guarantee of the value of a test. It is necessary to know now only how many persons pass an item but also which persons.

An item is appropriately diagnostic or has positive discriminating porer to the extent that it is passed by persons of ability greater than the level corresponding to the item and failed by those of ability less than the level corresponding to the item. An item passed equally often by persons at all levels of ability has no discriminative power. An item passed by relatively more persons of low ability than those of high ability has negative discriminative power; that is, while it discriminates, it does so in the wrong direction. Although a very hard or a very easy item cannot have a very high discriminative index for the whole range of ability, such an item may discriminate well over a narrow range. A very difficult item, for example, may be failed by all of the lower four-fiftho of a group of subjects, but if it is answered by as many as one quarter of the highest fifth of the group, it will be a vary valuable item.
$I_{\text {Adkins, }}$ p. 148 .
matie 6

ITEK DIFFIGULTY OF THE PRELTMHARY TEST, PART I: NUBBER OF ITEHS FOUND FOR EACH LEVEL OF DIFEICULTY REPRESENTED

Level of Difficulty
Number of Items
(Proportion of Total Group Pabsing)

| 2.7 | 1 |
| ---: | ---: |
| 5.5 | 4 |
| 8.3 | 3 |
| 11.1 | 8 |
| 13.2 | 77 |

13.3 11

| 16.6 | 5 |
| :--- | :--- |
| 19.4 | 7 |

22.2 7
$25.0 \quad 12$
27.78
30.5 6
33.3 8
36.1 3
30.8
4.6 $\frac{17}{4}$

山.4 8
47.2 6
50.0 7
52.7
58.3 6
61.1 3
63.8 I
66.6
69.4 I
72.2 I
75.0 I
80.5 I

Ideally, each test iten is mastured against or coxrelated with an independent or "extermal" criterion, just as the validuty of a total test score is deterrined by correlating it with an Independent measure of job performance. Hence item analysis is made against an external extterion nhenever possible. 1 .

When an external criterion ia lackings the items may be inproved by an analysis of their comelation with m "internal" oriterion, meaning the total scoro on whe tebt itself. Items thes analyzed are sometimes referred to as providing an incex of validity, meaning internal consistency, but this procedure in itsolf does not establish the validity of a test.

If the test as a whola does not have valicity, selection of items which correlate highly with the original test as a whole can never yield tt. What bhe process can accomplish is to select items whioh tend to measure whatever the test as a whole neasures. If the test is measuring what it is Intencied to mesure, then and then only will the least valid items be culled out by use of an internal. criterion. This process is more appropriately Viewed as a means of increasing the reliability (internal consistency) of a test per unit longth then as a method for ralidatimg a test. ${ }^{2}$

In ilien analysis using the total score of a test or subtest ag the internal criterion, it is necessary that the items be homogeneous; that it, that they measure "the sane psychological process or combination of processes. ${ }^{3}$ If items are correlstea against the score of the whole test, the assunption 18 that "all the items throughout the entire test are expected to be monogeneous in basle functions neasureci. 44
$I_{\text {Adkins, }} p \cdot 181$.
$2_{\text {Adkins, }} p \cdot 182$
$3_{\text {Freeman, }}$ p. 26.
$4_{\text {Ibid. }}$

In teste having subtests involving different areas of knowledge or different skills, it is advisable to set up separate internal criterion scores for each factor present in the total score. An example of this would be a test involving information and arithmetical problems. Since there are two distinct areas or psychological functions involved, it would be necessary to correlate the items with their respective subtests, and not with the total test scores. To do the latter could lead to inappropriate and misleading conclusions as to the relative validities of the scores.

The present test vas expected to involve only one function, $1 . e_{0}$ correlating the bounde of the harmonies of a musical passage with their corresponding notation. The homogeneity of the process has already been partially demonstrated in comection trith estimating the rellability of the test. To attempt to classify the items according to more than this one general process or function appeared to be both inpractical and unecessary.

There are many ways of detemining how well a single item agrees with the total score in the kind of discriminations it makes wong individuals. Most of these procedures are correlation methods. The correlation methods that have been used in this connection include the biserial ry, point biserial ry phi coefficient, and tetrachoric $r$. Whatever index of correlation is used, one variable is of necessity a dichotomy, as a rule. It is a division of the sample into those who pass the item and those who feil it. . The total test score is a continuous variable, but for convenience it is usually artificially dichotomized for the same of simplicity of numerical operations.

[^28]The choice of a statiztic from those mentioned depends "paritiy on the purpose for winch the test and the item analysis data are to be used, partly on the comonience with which each statistle serves that purpoee, and partily on the ease and economy of computation required by the practical circunstances. ${ }^{1}$

Somotimes it is deasred to seleot thoms for future two wich discriminate botween two extrene groups, such as the upper end lower ten: per cont of the population tested. Sometimes the boundaries are extended to include the highest and Iowest fourth, or the upper and lower fifty per cent. A comonly used method is to choose the upper and lower twenty- seven per cent. In this test the sample was thintromix subjecto, and 24 appeared advisable to take all the scores on oach iten into con sideration by dividing the group into the upper and lower fifty per cent.

A discussion of the advantages and disedvantages of each of the possible methods of analyzing the discriminating power of the items is not presented here, since they are thoroughly explained in the literature on the subject. The phi coefficient mas selected for tise because it offered certain advantages, was favored by several authorities; and seemed to fit the conditions.

If the oriterion is $a$ notural dichotomy and mast be used as such, an acceptable motkod of expressing the item-critarton relationship is the phi ccexficient when the group with which the test is to be used is essentially the sane with respect to level and distribution of ability as the group used for item analysis purposes and when the point of dichotomy in the criterion variable remains constant in successive groups in which the test is being used for predictive

$$
I_{\text {Devis, }} \mathrm{p}_{4} 289
$$

purposes. When . . these conditions are met, the phiz coerfieient is appropriate because it is a rigorous product-noment ir subject to precise teste of significance and sutbole for use in computing multiple regression weights. ${ }^{2}$

From several points of view, according to Guilford, experience has shown that the phi coefficient is among the most aatisfactory indices of item consistancy.

The computational aids . reduce the effort to a minimum which some othar mothods equal but do net go below. Its direct ielation to chi square provides automatically a test of minimal significant confficients for mat elize of sample. 2

If the number of examinees in ach of the diohotomous groups is the same, the phi conffiefents yay be read dinectly from a table prepared by Jurgenson. ${ }^{3}$ It ras possible to ase this tavle because every examinse that the opportuntty to constder and racis every itsit. Thus it was nocesm sary only to caleulete the value of phi only once for all the itoms, for any deatred level of confidence.

Phi. As derived from a $2 \times 2$ chi square by the equation 4

$$
\phi-\sqrt{\frac{x^{2}}{N}}
$$

It is customary to select the .01 or .05 levels of confidence as being significantiy difterent from chance. Therefore the values of . OL and .05, with one degree of freedom, were substituted in the fomula, as follows:
${ }^{1}$ Davis, pp. 290-291.
${ }^{2}$ Guilford, p. 500.
3. E. Jurgenson. "Table for Determining Phi Coeificients." Pbychometrika, 12 (1947) 17-29.
$4_{\text {Guilford, }}$ p. 340 .

$$
\begin{aligned}
& \phi_{.01} \sqrt{\frac{6.635}{36}}=.429 \\
& \phi_{.05}=\sqrt{\frac{3.841}{36}}=.327
\end{aligned}
$$

To determine borderline cases, phi was also computed at the .10 levele

$$
\phi_{.10}=\sqrt{\frac{2.706}{36}}=.274
$$

It was then necessary only to divide the test papers into the upper and lowro fifty per cent, and calculate the frequency of correct responses for each item, for each group. Jurgensen's table was entered horizontally with one derived proportion and vertically with the other. The proportions for each group, and the values found, are given in Table 60, Appendix 0.

## Results of Itom Analysis

Of a total of 150 items, 83 proved to be discriminating at the .01 or .05 levels of conflidence. In addition there were six items which just exceeded the .05 level and which could possibly be useful. The difficulties of the items were fairly well distributed throughout the possible ranges, whioh was desirable for the purpose of the test, but as a whole the test appeared to be too difficult. Although many of the too difficult items were to be excluded from the final form of the test, yet those difficult items which discriminated were too valuable to elininate. Also, some valuable items within excerpts had to be elfininated because the other items in the excerpt proved to have no discriminative value. It was therefore found necessary to obtain some easier items, as well as a few more discriminating items, by constructing another test along similar lines, and subjecting it to the same analysis as this form.

## The Preliminiary Test: Part 2

Selection of Additional Excerpts and Introduction of Errors

The analysis of Part 1 of the preltminary test had revealed a need for easier fitems, ie., errors in the performance of notes or chords Which could be nore easily detected by more musictanse Therefore nine new musical passages were selected from simple choral music. Into these passages logical exrors in performance were introduced. In addition, five excerpts from Part I were used again, but with some nev performance errors introduced, and each being recorded anew. The test consisted then of fourteen passages in which there were a total of forty-five errors (items).

Reproduction of the Scores

The musical scores were reproduced as before, legible manuscript on stencils. The form is included in Appendix $D_{2}$ with the notes performed incorrectivy marled in red symbols, for ease in referring to the subsequent tables of Item analyses.

Recording of the Test

The test was recorded in the same manner as Part $I_{9}$ using the same singers, and following the same general procedures. The master tape was made up of the best version of each excerpt, again with special reference to the clarity of the intended errors, but also with due regard for the best performance of each excerpt as a whole. An improvement over the use of a pitchpipe in sounding the starting pitch was to have the
pitch of the starting note or chord humed by the singers. This simulated actual rehearsal procedure, and aided in the orientation of the subjects to the tonality of each successive musical selection.

Interviews with mang of the persons taking Part 1 Indieated that the time Interval of ten seconds betmeen the annowncement of the number of an excerpt and the first periformence of the paseage did not need to We so long. It was therafore recuced to 隹ve geconds. Tum the complete procedure for one excerpt was es follows announctuent of the number of the excerpt; approxinately five seconds of tilence; the humed starting piteh follomed by a one-second interval of silonce; the Iirst perfomance of the excerpt; approximately five seconds of silence; the humed startm Ing pitch again followed by approximately one second of silence; the exact repetition of the excerpt. Approximately five seconds then intervened before the anomement of the number of the next excerpt.

The completed master tape mas listened to oritically by the same three capable mosicians, with particolar reference to the clarity of the intended errors. Again, there was agreement that the intended errore in the perfomance were all clearly evident as deviations from their corresponding notation.

## Warking and Scoring

The analysis of the mathods of marking and acoring the itoms in Part 1 resulted in the decision to use the simplest method, since the rosults were so similar, and since the simplest nothod provided ample discrimination anong the subjects in the sicill being tested. The subjects
were therefore instumeded to checi only the bowes above the notes or chords which were performed incorrectly: In scoring, each box correctly checked contributed one point fowerd the total score. No correction was made for chance success or guessing, but the subjects were not actived of thato.

## Questiomaire

No questiomaine was used rith this fom, since most of the subjects were the sme as for Part 1 , and since the questionnara mith Papt I hed vielded suffictent information to determine the form and ontent of the questomaire with wres to accompany the final sozm of the test.

## The Sample

A11 of the finct Eroup of thirbymit eubjecta covad not be obu thined to take Part 2. Homever, trenty-tiree of the Eromp ware the sman Andivicuals; and five more mere deam from the mame undergraduate and graduate lovels, matine a total of twentyelght musicsans. For all proctical purposes this group wres regarded as homogeneous with the first group of thirty-six, and the excerpts with their attendant items selected from this test could be included nith those of Part 3 . In arranging the Sinal form of the tert.

## Reliability

The reliability of this part of the prelininary test was estimated in order to ascertain whether it was as consistent in providing the same results as Part 1. As with Part 1, a product-moment $I$ based on the scores of split halves using alternate $\pm$ tens, and corrected to full length by the Speaman-Brom prophecy formia, was oalculated, using the data presented in Table 61, Appendix D. The coefficient of correlation of .94 which was obtained indicated that the test was very consistent in its resulte, especially since it contained only forty-five items.

## Analysis of Item Difficulty

Bech Iten was analyzed for difficulty, as in Part 1. The numer of items found at the various levels of difficulty are given in Table 7. The levels of difficulty were determined by the proportion of the total group of twenty-eight subjects who answered each item corractiy. The difficulty of each spocific item is given in Table 62, Appendix $D$, under the columin headed "per cent passing, total group."

The rance of difficulty of the test items approximated that of the first part of the preliminary test, but there were approximately as many easier itens as more difficult ones. The test had accomplished its purpose by providing a number of easier items.

## Analysis of Item Discrinination

The phi coefficient was used again in expressing the relationship of each item with the criterion of total scores. Using the formula

## TABLE 7

 ITMMS FOUN FOR DAGH LEVEL OT DTFFIGULTY REPRESENIED
Level of pofficulty
(Proportion of total group passing)

$$
\phi=\sqrt{\frac{x^{2}}{N}}
$$

the values of phi necessary to be significant at the .01, .05, and . 10 levels of confidence were calculated as follows:

$$
\begin{aligned}
& \phi_{.01}=\sqrt{\frac{6.635}{28}}=.487 \\
& \boldsymbol{\phi}_{.05}=\sqrt{\frac{3.847}{28}}=.370 \\
& \phi_{.10}=\sqrt{\frac{2.706}{28}}=.310
\end{aligned}
$$

The results of the analysis are given in Table 62, Appendix D. Twenty-two of the items were found to discriminate at the . 01 or .05 levels of confidence, whth one other item so close to the .05 level as to be constdered for inclusion in the final form of the test.

## Sumpery

From Part 1 and Part 2 of the preliminary test a sufficient nume ber of discriminating itens were found to construct a final form which was intemaliy consistent. These items also included a wide range of difficulty, so that it would be possible to discriminate among individuals as well as groups at many levels in the pover measured by the test. Final selection of the items, however, was also dependent upon other factors. These are discussed in the next section on the construction of the final form of the test.

## The Final Fom of the Test

The content of the final form of the test was dependent upon a number of factors besides that of including items of suitable difficulty
and discriminative porer. Among these were: (1) Practical time limits, (2) The inclusion of non-discriminating itemg (3) The arrangement of the musical excerptis according to the difficulty of the attendant items; (4) The inclusion of musical excerpts which were representatite of average chorel literature, and (5) The representativeness of chordal vocabulary as found in the test items.

In addition, the best method of reproducing the musical scores had to be decided, and the final recording had to be made.

## Practical Tine Limits

The use of more than one reel of magnetic tape was not desirable, because changing reels during a test period was time-consuning and distracting. (The exploratory test had required one thirty-minute reel and one fifteen-minute reel; Part 1 of the prelininary test had required two thirty-ninute reels, and only part 2 had required less than one thirtyminute reel.)

A test of such length that it could be administered during the usual class period was desirable. Such pertods usually range from 45 to 50 minutes in length. Part of the perlod would be taken in filling out the questionnaire which was to accompany the test, and in reading the test instructions. This was estimated to require from eight to ten minutes. Therefore the test should not be over 35 or 40 minutes in Iength.

The average commercial half-hour reel of tape, it was found, actually contained approximately thirty-two minutes of playing time.

It was further found that an additional four minutes of tape could be accomodated on the reel without creating mechanical difficulties or causing any noticeable distortion in the reproduction. A test of approxfmately thirty-six minutes' length was therefore decided upon. The masi ter tape and duplicates of the master tape which were subsequently made were extended to that length by splicing on the appropriate amount of extra tape.

The timing of the intervals between excerpts in Part 2 of the preliminary test seemed to be atisfactory according to opinjons obm tained from the exsminees, with the exception that the interval between the first pleying of an excerpt and its repetition, and between the repetition and the announcement of the next excerpty could just as well be reduced from five to approximately three seconds. Therefore, the total time which each excerpt would require mas calculated as follows: anouncement of the muber of the musical passage, I second; interval of silence in which to examine the score (during which the tape would continue to run), 5 seconds sounding of the mamed pitch of the note or chord on which the singers would start; 1 second; an interval of preparm ation to sing, such as would nomally occur in rehearsal or performance, 1 or 2 seconds; the musical passage (each one had to be timed individually); an interval in which to reflect on what had been heard and seen, 3 seconds; the humned pitch again, 1 second; another preparation for singing, 1 or 2 seconds; the repetition of the excerpt; and another interval before the anouncement of the next excerpt. 3 seconds. Thus for each excerpt approximately 16 to 10 seconds had to be allowed in addition to the time allowance for the passage itself, plus time for its repetition. Since
only a limited number of excerpts could be inciuded in a tape of thirtym six minutes duration, the relative merits of each excerpt had to be weighed carefully, taking all factors into consideration as much as possible.

## Inclusion of Non-Discriminating Items

Those items which did not discriminate between the high and low scoring groups in the prelumary tests at the on or . 05 levels of confidence vere excluded from the final form as far as possible. There was no problen in deciding vhether to accept or reject for use those excerpts which contained only one error (item). But in many instances an excerpt contained both discriminating and non-disoriminating items. This was alnost inevitable in a test of this type, for an underlying concept in construoting the test was that a lengtwy musical passage would be inefficient and time consuming if it contained only one error (item) in performance. In order to test the skill of coordinating ear and eye in detecting errors in performance in a simulated functional situationy in a test of reasonable length, it was necessary to introduce a number of errors in many of the passages. It was only to be expected that some of the errors mould not discriminate, but these errors could not, in many instances, be eliminated without affecting the context of notes or harmonies of adjecent errors which did discriminate. For the same reason some difficult items had to be retained. It was therefore necessary to consider carefully the merit of each excerpt with its attendant errors.

The oxecpts and items which were retained in tho find form are. diven in Table 63, Appondix E. Their source (Fart 1 or Part 2 of the
 are tabulated. Reference may be made to the printed fom of the test also included in Appondix Ey if it is desired to study the actual musion notation and context of the items.

The final form containod thiry musical excerpts having a total of 118 exrors (itens). Efehty-four of these atseriminated at the .01 or . OS level of confldence, five more were very close to the 05 level. Thege items ranged in difficulty ifrom easy to vory difficult, with inost of them spread throughont the midle ranges. Nith this number of satisSactory items it mas believed thet the test mould be eble to accompligh its purpose. The non-ciserilminating items could not have been eliminated whthout $a$ great deal of Inbor in remecording many of the excerpts. Some could not have been eliminated pithout changing the entire context of soma. To have done so mould have resulted in invalidating the entire excerpt. Therefore; with the few exceptions noted in the footnotes to Tablo 63, excerpts with their attendant items mere accepted or rejected Eor use in their entirety.

## Order of Items

Usually in constructing a test the oastest items are placed first, followed by tho of gradually increasing difficulty. It was not posmsible to follow this plan in this test because many of the excerpts contazned both easier and nore difficult iterns. It appeared to be more
psychologically sound to arrange the test so as to begin with easy bym tunes in the ustal four-part hamony (which incidentally contained some easier items), and progress to chorales and modal hamonizations which were all somemat similar in style. Theng in ordor to allay the inpression that the test was becoming more difficult, there followed a group of three-part chorale harmonizations. These were followed by selections from other choral 1iterature. Throughout, an attempt was made to keep the progression of ley changes with successive oxcerpts as smooth as possible. With the time intervals and announcement of mubers between the musical passages; the effect of abrupt ohnges in leey; and hence possible loss of orientation, was kept at a minimum.

## Hepresentativeness of Choral Husic <br> Included in the Finel Form

Since the test was in the choral medium it would probably be of particular Interest and use to prospective and active choral directors. Therefore an effort mas to have representative samples of the musical literature which the average school or charch choir might use inm cluded in the final form. To this end a record was kept of all the various samples which had been used in the preliminary tests. The number which could be included in the final fom mas, of course, severely limIted by considerations of time limits for the test, and by the necessity of having each sample include errors of performance which discriminated.

How well the effort succeeded in including a representative sample of reedily available music whioh the average choir might attempt may be judged by examining Table 8. In any event the test was considered to
include an adequate sampling of standard musical literature of this type. Representativeness of Chordal Vocabulary

Since the skill measured by the test is presumably developed primarily by training in harnony and ear-training, the structure of the melodic and harmonic intervals and chords making up the test items would possibly be of interest to teachers of those subjects. It was believed that the errors in performance should occur in a representative sample of at least the commonly used chords.

The intervals and chords into which errors of performance were introduced were analyzed according to their score notation and tabulated for the exploratoxy test and for each part of the preliminary test. This analysis was made on the basis of the quality or mode of the interval or chord, such as major or minor, and on the basis of function, as that of a chord being $I$, $I V_{p}$ or $V_{\text {, etc. . As far as other considerations }}$ affecting the content of the fingl test form would allow, such as whether the test errors (iteme) discriminated or whether they were of suitable difficulty, the errors were chosen for their contribution to an adequate sampling of the comnonly used chords. More incidentally, they were choson for their contribution to a sampling of some of the conmon melodic and harmonic intervals.

The chords and intervals did not by any means represent all the possible structures and functions, but many of such possibilities occur but rarely. In chromatic harmony and modem harmony such possibilities occur more often, but, as stated earlier in the study, the music was

## TABLI 8

OHORAL MUSIO REPEESEMTED TH THE FTMAL TEST FORM

| Excerpt | Style | Wationality | Century | Parts | Staves |
| :---: | :---: | :---: | :---: | :---: | :---: |
| 1 | Hymin | Enclish | 20 | SATB | 2 |
| 2 | Hivm | American | 20 | SATB | 2 |
| 3 | Hymu | German | 18 | SATE | 2 |
| 4 | Chorale | German | 16 | SATB | 2 |
| 5 | Hymin | American | 20 | SATB | 2 |
| 6 | Chorale | Cerman | 18 | SATB | 2 |
| 7 | Hymn | German | 19 | SATB | 2 |
| 8 | Psalm-Tune | American | 20 | SATB | 2 |
| 9. | Chorale | German | 18 | SATB | 2 |
| 10 | Chorale | German | 18 | SATB | 2 |
| 11 | Secular | American | 18 | SATB | 2 |
| 12 | Chorale | German | 18 | SSA | 2 |
| 13 | Chorale | German | 17 | SSA | 2 |
| 14 | Chorale | German | 18 | SSA | 2 |
| 15 | Chorale | German | 16 | SSA | 2 |
| 16 | Chorale | German | 16 | SSA | 2 |
| 17 | Chorale | German | 16 | SSA | 2 |
| 18 | Motet | German | 16 | ATB | 2 |
| 19 | Anchem | Italian | 16 | SATB | 2 |
| 20 | Anthem | Russian | 19 | SATH | 2 |
| 21 | Anthem | Russian | 19 | SATB | 2 |
| 22 | Psalmbinme | American | 20 | SATB | 2 |
| 23 | Secular | American | 18 | SATB | 2 |
| 24 | Anthem | Italian | 19 | SATB | 2 |
| 25 | Madrigal | Italian | 16 | SATB | 3 |
| 26 | Motet | French | 74 | SATB | 3 |
| 27 | Secular | American | 20 | SB | 2 |
| 28 | Secular | American | 20 | SB | 2 |
| 29 | Secular | Anerican | 20 | SATB | 2 |
| 30 | Secular | American | 20 | SSATB | 2 |

selected on the basis of tits liselihood of use by the average choir. This restrioted the choice mosily to music of a diatonic style, with which nore choirs are familiar, and in particular the masic based on the harmonic practice of the eleghteenth and nineteenth centuries.

Just as this study was concerned with ineasuring a skill found useful in a practical situation, so the concern with the representative ness of the chovdal vocabulary was not with the theoretical logical possibilities of chord structure, as found in many harmony textbooks, but with the frequency of harmonies as they actually occur in music. Therefore the representativeness of the chords sampled must be judged with the aid of some knowledge of actual chord frequencies. Such knowledge has been supplied in a study by Budge. ${ }^{7}$ Through an analysis of the music of representative composers of the oighteenth and nineteenth conturies, she found the order of frequencies of diatonic chords to be as follows:

| $1-1$ | 12-IV6 | 23-VII | $34-$ VII 6at |
| :---: | :---: | :---: | :---: |
| 2-V7 | 13-74-2 | 24-III 6 | 35-III 6-4 |
| 3-V | $14_{4}-\mathrm{V} 4-3$ | $25-\mathrm{V} 6 \mathrm{~m}$ | $36-$ WI $6-5$ |
| $4-16$ | 15-1V6-4 | 26- II 4-2 | 37-17 |
| $5-$ I 6-4 | 16-VII 6 | 27-IV 7 | 38-VII 4-2 |
| 6-IV | 17-II 6-5 | 20-VI 7 | 39- VI 4-2 |
| 7-VI | 18-WI 6 | 29-VI 6-4 | 40 - IV 6.5 |
| 8-II 6 | 19-III | 30-VII 6-5 | 41 - III 7 |
| 9-76 | 20-II 7 | 31 - II $4-3$ | 42 - IV 4-2 |
| 10-II | $21-79$ | 32 - VII $4-3$ | $43-$ IV 4-3 |
| $11-$ V $6-5$ | 22-VII 7 | $33-$ II 6-4 | $4 L_{1}-$ VI $4-3$ |

I Helen Budge. A Study of Chord Frequencies Based on the Husic of Mepresentative Composers of the Mjehteenth and Jineteenth Centuries. Bureau of Publications, Teachers College, Columbia University, Mew York, 1943.

Chorda used once each were: I 4-2, II 9, III 9, IV 9, III 6-5; III 4-2. Chords never used were I 6m5, I $\operatorname{lu} \mathbf{- 3}$, I 9, III 4-3; VI 9; VII 9. Chords which ranked below five were used lesg than 5 多 each. Chords which ranked below 9 mere used less than at each. Chords minch ranted below 16 wexe used less then ly each.
 Chords which ranied belov 32 were used less than $1 / 10$ of $3 \%$ each.

From this Budge derived the generaltuation that the following chords are used less than $\frac{7}{2}$ of $1 \%$ each:
2. All stivfour chords except the I6-lt and IV6-4e
2. A11 fundemental seventh chords except the II7 and V7.
3. A11 inverted serenth chords except the $I T 6-5, V 6-5, V 4, j$, and $74-2$.

In Table 64, Appendix $H_{2}$ is presented the analysis of the quality (or mode) and function of each best item (chord or intervel) retained in the final test formy as it appeared in the notation. Table 9 provides a sumary of these data. From these tables, keeping the results of Eudge's findings in mind, the sampling of the chordal vocabrilary represented by the test items alone was judged to be satisfactory. Put of course the entire context of notes and chords hac to be considered by the subjects in order to dotomine wherein the exrors of porformance occurred, and obviously thirty excerpts, wo of considerable length, would contian a great vartety of melodic and hamonic intervals and chords. And since the actual errors occurred in a large sampling of chords and intervals, it appeared to be unnecessary to analyze overy interval and chord in order to demonstrate the representativeness of chordal vocabulary included in the test.

TABLB 9

SUMARY DATA OF ANAEISES OF THE QUALITY AND FUNCTION OF THE CHORDS




[^29]
## Reproduction of the kusical Scores

Since more copies of the final form of the printed test were rem quired than could be reproduced legibly by stencils, the photo-offset method was utilized. By this method, which is used widely not only in many types of tests but also in much commercial printing, it was possible to reproduce any deaired number of copies of manuscript as legible as the original itself. In order to reduce the number of pages required, the notation of the musical selections was arranged on five sheets of legal size paper ( $8 \frac{1}{2} \times \mathrm{I}_{4}$ inches), which appeared to be a convenient size to handle.

One alteration was made in the appearance of the scores. This was to place the small blanks or boxes, in which checks for errors were to be marked, in a straight line below the bass staff, with one blank vertically under each note or chord. In the preliminary tests they had been placed in a horizontal line above the treble staff, but the new arrangement permitted a quicker check mark to be made, and without the possibility of having the hand obscure the notation while placing a check marls.

Experience with the preliminary tests had shown that the examinees had no difficulty in understanding the purpose and procedure of the test. Therefore, only one example was included in the final form.

Hecording of the Test

The final tape recording was made up from mong all the musical passages which had previously been used, and from the few passages which
were remecorded. The process of making the master tape by using two machines, and playing the desired versions through the first machine into the recording input of the second, has been described previously. In order to reduce possible loss of fidelity to a minimum, the original recordings of versions selected for use were used, rather than taking them from the master tapes of parts 1 and 2 of the preliminary test. In addition to the excerpts, the announcement of numbers and the humed, starting pitches were recorded and "dubbed" onto the master tape in proper timing sequence.

The possible loss in recording fidelity by this process had prem viously been judged to be unoticeable, and the intended errors in perfomance had also been judged to be clearly evident. From the completed master tape a number of copies were made, but even this process did not appear to cause any loss of overtones or other noticeable distoxtion in the sound. Three capable musicians could detect no difference in quality between the naster tape and the duplicates of it.

Test Instructions

The test inatructions for the final form were based upon those used in the preliminary test. These, together with suggestions from several of the musicians who served as subjects, were utilized in forming a brief but comprehensive statement of objectives and procedures. A oopy of the instructions is included for reference in Appendix E.

In addition, special instructions for the test administrator were prepared. A copy of these instructions is also included in the same appendix.

## THE DESIGN OF THE RXPHRTHENTAL STUNY

## Introduction

The first objective of the study was to construct an instrument to measure the achievement of musicians in the recognition of discrepancies between the notation and the performance of the tones and harmonies of choral music. The process of constructing the instrument has been described in Chapter III. There remained the further steps of estabu lishing the reliability and validity of the final form of the test.

A necessary procedure in the study, therefore, wes to adminiater the test to a sufficiently large sample of the groups for which it was intended, undergraduate and graduate music majors in selected collegea and universities.

The test was also administered to a eriterion group of musicians rated as being "above average" musicians. This was done to establish the validity of the instrument by testing the underlying assumption that superior musicians possessed in significantly greater amount the skill to be measured by the test.

The aecond objective of the study was to detemine the relationship between selected factors or characteristics, and the ability to be measured by the test. Levels of achievement were sought for the factor thought to contribute most to skill in reading and hearing the scorem-the anount of training in music theory. In order to obtain the data necesm sary for the study, a questionnaire was devised and administered together with the test.

## The Questionnaire

In Chapter III the formulation of a prelininary form of the ques tionnaire was described. This was analyed for inprovernents and addtions, and subjected to criticisme by three college and university music instructors, until the final form energed. A copy of the questionnaire is included in Appendix. E. This was kept as brief and as simple to answer as possible, since the aciministration of the test in its entirety, including the questionnaire, reading of instructions, and the test prom cedure itself, was to be accomplished within the time limft of the usual 45 on 50 -minute class period. However; most of the factors winich were thought to have some relation to achievement on the tost were included.

The most evident of these was the amount of troining in music theory, since this is directly related to understanding the elements of music, their structure and function. Accordingly, one question was dew voted to identifying those usual courses in theory which had been completed and those witich were in progress. The purpose of the double information, as well as a question concerming the present school level, Was to pinpoint the exact training status of the individual.

Other characteristics included the following: sex, age, principal performing medium, number of years of atudy and experience in that medim, number of years of atudy and experience on piano (non-piano majors), piano playing ability (ability or inability to play hym tunes and simple accompaniments at sight), piano improvising ability (ability or inability to improvise some type of piano accompaninent for simple melodies), the part usually sung in choral ensemble, amount of choral directing
experience, and anount of instrumental directing experience.
One question was devoted to most recent school marlis obtained in those theory courses which are comon to the first two years of the college masic curriculum-harnony, ear-trainings sight singing, and keyboard harmony, or thoir combination, "theory".

Despite the known ureliability of school marks when used as criteria with which to compare achiovenent as measured by something else, It was believed that a comparison of marks. pith achievenent on the test would be weful. If marics in those courses most closely related to the development of the skill measured by the test corresponded with test achievement; an additional indication of the validity of the test would be provided.

School marks generally fall into five categories equivalent to superior, above average, average, below average, and failing. Since the test was to be administered to some mustions who were several years romoved from their most recent course of those specified, they might not remember specific grades, but would probably remember their general level of achievenent as being above average, average, or below average. The instructions, therefore, were to indicate the appropriate category. Specifically, maries of "A" or "B" or their equivalent were to be considered as "above average", "g" or its equivalent as "average", and "D" or "TF" or their equivalent as "below average".

The final question, to be answered after the administration of the test, was concerned with two parts: (a) the method used predominantly in listening for accuracy of harmonic effect, and (b) the extent of chord
analysis or rocogntion of their function as boing tonic, dominant, etc. Part $a$, concemed with accuracy of hamonto effect, was morded in such a Way as to detarmine which of the folloming athods was used predominantiy. Interrogation of many subjects cooperating in the preliminary test had indicated that the principal nethods weres (1) to listen to one or two parts at a time, i.e., nelodically, (2) to listen to the quality of each chord as an entity, 1.e., the determination of the nodo as being major, minnor, etc.g and (3) to 1isten for the general appropriateness of harmonic and melodec effect and style, i.e., without molodic or ohordal analysis. Since these methods are not mutually axclusive, everyone probably using each one to some extent, the question was phrased so as to determine only which rethod was used predominantly. Part b provided four categories for the extent of anelysis or recognition of chord function: (1) most of the time, (2) irequently, (3) occasionally, and (4) never. The entire question was included because or the probrble value of the derived information to theory teachers.

## The Sample

A lotter explaining briefly the nature and purpose of the test, and solioiting cooporation in the administration of the test, was gent to the heads of musle, music education, or thoory departnents of twenty selected unlversitios, colleges, and conservatories. An effort was made to have an adequate representation arong the schools as to Jocation and type of sohool, but a large and representative random sampling was beyond the scope of the study, which was experimental in nature.

Of the twenty schools, eighteen wore interested in the study, but only ter were able to tako sufficient time from their schedules to administer the test. of these, eaght completed the adrinistration and returned the matorials, A total of 265 tests were returned from these eight sohools, but five had to be discarded because of very incomplete questionnsire data, leaving a total of 260 tests.

The type of school, and number of tests recsived from each, are given in Table 10. These schools were located in five statest Kansas, Missouri, Colorado, 0klanoma, and Tennessee. The maber of tests mas regarded as sufficiont for statistically reliable results, because the samples mere to be pooled if possible; and no comparisons between schools were to be made in this study.

The Criterion Group

The validity of the test, as stated previously, was based upon the assumption that a group of musicians rated as being above average would possess more of the characteristics involved in the total complex of mastelanghip, and to a groater degree, than would average or belowaverage musicians such as mould be found in the unselected group provided by the sample from the schoole. Anong these characteristics would be the possession to a greater degree of the ability to be mexsured by the test.

There was no attempt to deternine the superior or "above-average" musicians by jucement of a panel of experts using some type of rating scale. On the contraxy, it was believed that the validity of the test would be better established if a mixed group of musicians, rated as

## TabIM 10

SCHOOLS PARTICTPATING IN THP STUDY, AND WUMBER OE TEST RBSULTS REOSIVED FHOE EACH

| School | Type | Number of Tosts |
| :---: | :---: | :---: |
| A | State Oniversity | 96 |
| B | State College | 52 |
| 0 | State University | 25 |
| D | Jmior College | 32 |
| E | State College | 19 |
| F | Teachers College | 11 |
| Q | State University | $\mathrm{IH}_{4}$ |
| H | Liberal Arts College | 11 |
|  | $N=260$ |  |

being botter than average ky a Per ainple criteria of masoal succoss, succeded in scoring slanifiowity hfgor on the tert than a group of unselecter mosteiang, such as would be forndin a sample of college studont ausicisns. Thus, two general critowis were applied: (I) occupation of an edranced position of tusical responsibility; (2) achiovement of superion mosical resulte In the first could be placed hends of music depentments in colleges and untersithes, theory teachers, gustc adjudicatorgy stc. Tn the second could be plonen high schonk and college Instrunentol and choral directore who produced outstandine resulta with their groupss and equpeserg whth pullished works. The tro aritexha were not, of course, mutualy exolusive.

Since cooperation in the study by mombers of the criterion group involved personal prestige, dempte assurancen of mongrity, it was difificult to secure a large muber. 3owerex, mentymive masicians to whom either or both of the catteria applied were individually secured, and this number was regarced as sufficient for this phase of the stucy. The range and nature of thes group, and the gereral oritoria by which they wore selected, are detailed in Table 11.

## Statistical Tools Used

For comparisons involving status according to the arount of train Ing in theory, the data from the various schcols could be pooled if thsy were homogeneous. The chi square stathistic mas used to test the hypothesie of normality of distribution of the scores.

Oht square was also used in sturfing the differences in scores between the subjects classified according to aach factor or characteristic.

## TABLIS 12

## HEMBERS OF THE CAITERION GROUP, AND GUALIFICATIOMS FOR HEGBURSHIP ACCORONO TO THE GEMERAL CRITERIA

| Member | Position or Occupation or other Hvidence of Superior |
| :---: | :---: |
| Lasicianship |  |

For example, age was one factor investigated.
Variation by the method of ranks was employed in deternining whether there was any relationship amons the scores of the critarion group and the scores of the groups classified according to the amount of theory trainfing. This was one of the methods used in validating the test.

T-tests were calculatea in testing the hypothesis that the mean of the criterion group was not significantly different from that of each of the groups classified according to amount of theory training. The assumption which had to be satisfied before the t-test could be applied was that there was no difference between the wariances of the groups compared. This assumption of homogeneity of variances was tested by application of the F-test. In the event thet this assumption was not sathsfied, the Behrons-fisher testy or d-test, was used in place of the t-test.

I-tests, d-tests, and f-tests were also used in comparing the means and varlances of the scores of groups classified according to marks in musie theory coursesy and in determining the reliablitity of subjects' information as to those marks.

The reliability coeffiolent of the test was computed by the produetmoment method, as with the preliminary forms of the test.

The results of the various analyses conducted were judged to be significant or insignificant according to a commonly used basis of judgment:

1. The results are said to be significant if the conclusion that they are would be orroneous in 1 per cent or less of the cases.

## 134

2. The results may be signtficant but further observations are necessary (that is, we suspend judgment) if the conclusion that the results are signiticant would be wrong in 5 per cent or less but more than 1 per cent of the cases.
3. The results are not signiffeant if our conclusion that they are slgmifleant would be in exror in more than:: 5 per cent of cases. ${ }^{1}$

IPainer 0. Johnson. Statistical Hethods in Pesearch. PrenticeHall, Inc.s New Tork, 1949. P. 28.

## GHAPTER V

## RESULTS AND IUTERPRETATIONS

Before reporting the results, it is necessary to call attention, in the section which follows, to the uses made of the questionnaire data, and to describe in cetall the manner of classifying the data.

Subjects Classified According to mount of Theory Training

Since theory tratning was thought to contribute to the development of the skill measured by the test, the subjects were classified according to the anount of training in theory. The data for this were provided by question No. 5 of the questionnaire, theory courses completed and in progress. Although question No. 4 concerned present school level, this was principally to provide an additional check on the accuracy of the response to No. 5. School level alone could not Indicate music training status accurately, since persons majoring in music do not necessarily always enter the masic curriculum in their first year in college. All the tests were administered toward the close of the spring semestery therefore the subjects were considered to have completed the course in which they were emrolled at the time of testing.

In question No. 4 were listed the theory courses found in comon In the music curricula of many schools. Of these, the ones that are basically concerned with reading, hearing and understanding the elements of music-malodio and harmonic intervals, chords, and rivtims-are harmony, eaf-training, sight singing, and keyboard harmony, or their

Integration into a general course in "theory". The content of these courses is fairly well standardized, and usually covers two gears' work. However, the contert, methods, and objectives of the more advanced courses in counterpoint, form and analysis, arranging, and others, are not as standardized. Therefore, the most logical classification appeared to be as follows: (1) one year or less of theory (2) two years or less, but more than one year, of theory, and (3) three years or more of theory. For corvenience, the categorles are henceforth labeled as one, two, or threemor-more years of theory training.

High school harmony or theory was not taken into account in classifying the subjects, since only a comparatively small number of subjects had had that course. Also, the content of that course is relatively unstandardized and, in any event, is not as comprehensive as the first yoar of college-level theory.

Conducting was also not taken into account in deternining the classification, since it appored that conducting was offered even on the first or second year level in sone schools. Conducting, furthemore, is often prinarily a course in anm movements and hand signals, with analysis of the technical content of music playing almost no part. All other courses above second year theory placed that person in the third year category, providing that sirst and second year theory had both already been taken.

Table 12 sumarizes the data obtained by the classification of the subjects according to the amount of theory training. Schools $G$ and F did not provide data for any subjects in the first two classifications,

SUBJEOTS CLASSIFIED ACCORDIWG TO AOOWT OF THEORY TRAMUING

| School. | 1. year | 2 years | 3 or more years |
| :---: | :---: | :---: | :---: |
| A | 29 | 47 | 20 |
| B | 31 | 10 | 11 |
| 0 | $\cdots$ | - | 25 |
| D | 19 | 13 | $\pm$ |
| E | 19 | - | $\cdots$ |
| F | - | - | 11 |
| c | 7 | 6 | 1 |
| H | 5 | 5 | 1 |
|  | 110 | 82 | 69 |

and School D (a junior college) had no subjects in the third classificam tion. School E administered the test only to first year students. In order to deal with the data as a whole, and to have a sufficient number of aubjects in each classification, it was necessary to determine whether the data for each classiflcation came from homogeneous populations and could be pooled.

Pooling of Scores of Subjects in Each Category of Theory Training

Chit square was the statistical tool applied to test the hypothesis that there was no difference in distributions of test scores between schools in each classiffication of amount of theory training. If the hypothesis could be accepted, pooling was accomplished.

In the analysis of the distribution of scores of subjects having one year of theory training, schools Ay By D and $E$ were included, $G$ and I 1 being excluded because of the small number of cases. The results of the analysis are given in Table 13. In this and the following tables of chi square, cells containing small theoretical frequencies ware combined whenever practicable in order to heve at least the minimum linit of five which Guilford suggests. 1 Thus, the small tail frequencies at each and were combined to obtain the cell frequencies given in the tables. In some cases the theoretical frequency of certain cells is much less than

[^30]
## TABLE 13




| Score interval | $\wedge$ | B | Schools $D$ | E | Totals |
| :---: | :---: | :---: | :---: | :---: | :---: |
| 30 or more | 14 | 9 | 5 | 6 | 34 |
| 20-29 | 9 | 14 | 7 | 4 | 34 |
| 19 or less | 6 | 8 | 7 | 9 | 30 |
| Totals | 29 | 31 | 19 | 19 | 98 |

$$
\begin{aligned}
& \chi^{2}=\sum\left[\frac{\left(f_{0}-f_{e}\right)^{2}}{f_{e}}\right]^{1} \\
& X^{2}=7.551 \\
& \text { d.f. }=6 \quad .30>P>.20
\end{aligned}
$$

IGuilford, p. 276. In the solution of cht square in this and subsequent tables, fe was calculated from the marginal totals.
five. The effect of small expected frequencies is often to increase the size of chi equere. ${ }^{1}$ Therefore, in such casea the null hypothesis of no difference in the distribution may be acepted with increased confidence if chi square is still not large enough to indicate a significant difference.

The obtained value of chl square was small enough to indicate that the differences in the distributions were due to chance. Therefore the smples could be pooled, since the distributions were not significantly different. By inspection, it was decided that the scores of the subjects in schools $G$ and $H$ could be combined with those of this group, maiting a total of 110.

The scores of the subjects from schools $A, B$, and $D$ were used in comparing the distributions in the clessification of two years of theory training. Table I4 gives the results. The obtained value of chi square was small enough to indicate that the null hypothesis of no difference in the distributions should be accepted, and that therefore the scores could be combined. inspection, the scores of the subjects in this classification from schools $Q$ and $H$ were included, making a total of 81.

Table 15 gives the results of the comparison of the distribution of scores in the classification of three or more years of theory train ing, using the data from schools $A, B, C$, and $F$. Again, the obtained value of chi equare was mall enough to indicate that the differences in the distributions were due to chance, and the scores could be combined.
$I_{\text {Guilford, p. }} 284$.

## TABLE 14

 HAVIMG TWO YEARS OF THEORX TPAIMING: $5 C H O O L S A, B, A N D D$

| Score <br> interval | A | B | D | Totals |
| :---: | :---: | :---: | :---: | :---: |
| 63 or more | 5 | 0 | 1 | 6 |
| $47-62$ | 9 | 3 | 1 | 13 |
| $31-46$ | 16 | 3 | 3 | 22 |
| 30 or less | 17 | 4 | 8 | 29 |
| Totals | 47 | 10 | 13 | 70 |

$$
\begin{aligned}
& x^{2}=4.323 \\
& \text { d.f. }=6 \quad .70>P>.50
\end{aligned}
$$

TABLE 15
 HAVIMG THEEE TEARS OR NOEE OE THEOKY TKATMIHG:

SCHOOLS $A, B, C$, AND $F$

| Score <br> interval | $A$ | B Schools | F | Totals |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
| $76-102$ | 1 | 1 | 6 | 2 | 10 |
| $51-75$ | 4 | 2 | 3 | 0 | 9 |
| $26-50$ | 10 | 6 | 12 | 8 | 36 |
| $8-25$ | 5 | 2 | 4 | 1 | 12 |
| Totals | 20 | 11 | 25 | 11 | 67 |

$$
\begin{gathered}
x^{2}=7.405 \\
\text { u.f. }=9 \quad .70>P>.50
\end{gathered}
$$

By inspection, the scores of the subjects in this category from schools $Q$ and 1 l wers included, making a total of 69.

The distributions of scores of the subjects in each category of theory trating were found to be homogeneous anong the schools included in the smple, end were pooled in the subsequent tests of significance betreen the neans of the categories. Before proceeding with these comparisons, however, the reliability of the test mas calculated.

## Roliability of tho pest

In the prelininary tests the reliability was calculated by correlating the split halves based on alternate items. For reasons preWhously stated this appeared to be the bost manor of dividing the test into presumably equal halves on a chence basis. By the nature of the test, it was a rather rigorous test of rellability.

In estimating the reliability of the test, a random sample of filety cases was decided upon, but the variable of amount of theory trainang that believed to be important, and therefore a stratified random auple was choden. Ench group contributed in proportion to its number to the total of fifty, The criterion group was also included, since thit group sas also to be msed in dotermining the validity of the test. Table 16 gives the proportion of cases contritbuted by each group to the total sample of ilifty. From each group a rantom sample was dram, using
table 16

## PROFORTION OF CASES CONTRIBUTED BY GROUPS TO A STRATIFTED SAMPLE OF FIFTY CASES

| Group | Munber | Proportion | Number of <br> Cases |
| :--- | :---: | :---: | :---: |
| I year of theory <br> 2 years of theory | 110 | .39 | 19 |

a table of rendom nubers. Table 17 Hists the cases drawn from each group.

By means of scoring keys the mumber of correct odd and even items in each test was computed. The data obtained are given in Table 65 , Appendx E. The cocricicient of correlation between the split halves was computed by the product-moment method, and the reliability of the fullJength test was estimated by use of the Speaman-Brown formule. The ooefficient of .98 which wes obtained Indicated that the test was highly reliable, or measured very constistently

The standard error of measurement was computed by the formulat

$$
\sigma_{t_{\infty}}=\sigma_{t} \sqrt{1-r_{t t}}
$$

Where $\sigma_{t}$ etemdard deviation of the distribution of obtained scores.
$r_{t \in}=$ reliability coefficient.
Qompubing $\sigma$ by the formala for original measurements, $\sigma=\frac{1}{N} \sqrt{N X^{2}-(X)^{2}}$ the standard exror was computedy $\sigma_{t_{\infty}}=24.33 \sqrt{1-.98}=3.41$. Assuming that the scores on the tegt are true scores, if the test were ropeated an infinite number of tines each obtained score could be

Helen H. Walker and Joseph Lev. Statistical Inference. Henry Holt and Co., Nev York, 1953. Pp. 484-485: a table of random numbers taken from a 30 mpege table of 105,000 random digits prepared by the Bureau of Transport Economics and Statistics of the Interstate Commerce Comission, washington, D. C. The numbers were yead as folloust for the first year group, colum 3, first three digitsp for the second year group, colwm 1 , firet two digits; for the group having three years of theory, colum 6, first two digits; for the criterion group, colum 5, fírst tro digits.

TABLE 17.

IMDIVIDUAL CASES FROM GROUPS IA A STRATIETED RANDON SAMPLE OF FIFTY FOA USE IN CALCULATHM TEST RELABILITY

| One year of theory |  | two years of theory |  | three or more years of theory |  | Oriterion |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 18 | (B17) | 23 | (A67) | 44 | (c16) | 6 |
| 105 | (a3) | 5 | (A93) | 43 | (c14) | 2 |
| 110 | (18) | 14 | (A91) | 34 | (c7) | 8 |
| 84 | (E14) | 38 | (A22) | 13 | (A78) | 17 |
|  | (D23) | 11. | (A79) | 50 | (c19) | 20 |
| 107 | (H2) | 43 | (490) | 36 | (c17) |  |
| 9 | (B9) | 49 | (B18) | 27 | (333) |  |
| 50 | (A51) | 36 | ( $\mathrm{A} 5^{\prime} 3$ ) | 20 | (A8) |  |
| 23 | (B5]) | 7 | (A83) | 47 | (c23) |  |
| 7 | (B4) | 61 | (DI) | 19 | (A80) |  |
|  |  |  |  | 54 | (010) |  |
| 74 | (D13) | 31. | (A27) | 69 | (G7) |  |
| 45 | (A1) | 57 | (85) |  |  |  |
| 34 | (A57) | 9 | (A21) |  |  |  |
| 43 | (A86) | 72 | (171) |  |  |  |
| 104 | (a13) |  |  |  |  |  |
| 33 | (A64) |  |  |  |  |  |
| 82 | (E10) |  |  |  |  |  |
| 15 | (B46) |  |  |  |  |  |
| 32 | (A82) |  |  |  |  |  |

Fote: The pooled groups in each category were first ranked from high to low by echools, and then the cases were numbered consecutively throughout each category. The letters and numbers in parenthesis refer to the schools and subjects as listed in the tables of original data in Appendix A. The criterion group meintains its original numbering.
expected to fall within the limits plus or minus 3.47 units two-thirds of the time. For example, if an individual received a score of 50 on the test, and he repeated the test an infinite number of times, without his score being affected by such factors as practice or fatigue, his obtained scores would fall botween 46.59 and 53.21 about two-thirds of the time.

## Comparisons Between Groups Classifled According to Anount of Theory Training

In order to determine whether the amount of theory training made a difference in achievement on the test, t-testis ${ }^{2}$ or d-tests ${ }^{2}$ of the significance of the difference of the means betweon groups classified according to the amount of theory training were applied. The tests were preceded by $F$ tests ${ }^{3}$ for homogeneity of variance. The results of the comparisons are given in Table 18. The varlances of the groups having one, and two years of training closely approached a significant difference at the . 01 level of confidence, and therefore the d-test, for determining the significance of the difference between means when variances are unequaly was applied.

It was found that the means between the groups having one and two years of theory training differed significantiv, but thore was no significant difference botween the means of the groups having two years,

[^31]TADIE 1.8

COMPARISONS BETHELH GROUPS CLASSIFIED ACCORDTHG TO AMOTNT OF THEORE TRATMING: t- and d-TESTS

| Statistic | 1 year of theory | 2 years of theory |
| :---: | :---: | :---: |
| H | 110 | 81 |
| $\sum X$ | 3,072 | 3,154 |
| $\Sigma x^{2}$ | 107.658 | 348,616 |
| $\bar{X}$ | 27.93 | 38.94 |
| $F$ | 1.6079 |  |
|  | d.f. $=80,109$ | . $05>P \gg .01$ |
| d | 1.477 |  |
|  | $n=81,210$ | $\mathrm{P}<.01$ |
|  | 2 years of theory | 3 or more years of theory |
| N | 81 | 69 |
| $\Sigma \pm$ | 3,154 | 3,060 |
| $\Sigma x^{2}$ | 148,616 | 162,868 |
| $\overline{\mathrm{X}}$ | 38.94 | 4.34 |
| F | 1.2384 |  |
|  | d.E* $=68,80$ | $\mathrm{P}>.05$ |
| t | . 2859 |  |
|  | d.f. $=148$ | $.8>P>.7$ |

and three or more years of traning. In order to tost this Important finding further, the homogeneity of disbribution of the scores of these two groups bas tosted by use of the che gquaro statistic, as shom in Table 19. The distribution was fomd to bo quite homogeneous.

These rosults were interpreted to mean that having one or two years of theory training made a significant difference in achievement on the test, but that training beyond that did not. Thus the decision not to attempt to subdivide the group having three or more yoars of training Into sub-crouns of three, fours five, of nore years of training was justified. The fact that there was a signifiont difforence in test achierenent between the first and second yoar groups, but not betmeen the second year group and the advanced group, Was attributed to the fact that the courses in theory furing the firnt two years deal nore specifically vith the understanding and recognition, both by sight and hoaring, of the structure and function of intervals and chords. Apparentiy the nore advanced courges do not contaibute significantly to the akill measured by the test, Other factors, guch as recency of training, may contribite to the fincing; but on the other hand, one would nomally expect the skill to improve with maturation end increased experience with music.

Validity of the Test

The usual coefficient of validity computed as a product-moment coefficient of correlation was not possible in this study because there was no outside criterion to provide numerical scores.

## TABLE 19

ANALTSIS OF HOICGBNEIIT OF DISTRIBUTION OF SCORES OF GROUPS HAVING TNO YEADS, OF THEORY, AND THRES OR YORE PEARS

| Score <br> interval | Tho years <br> of theory | Three or more <br> years of theory | Totals |
| :---: | :---: | :---: | :---: |
| 80.plus | 3 | 7 | 10 |
| $60-79$ | 7 | 5 | 12 |
| $40-59$ | 21 | 23 | 44 |
| $30-39$ | 19 | 16 | 35 |
| $20-29$ | 26 | 4 | 40 |
| 19 and less | 5 | 69 | 9 |
| Totals | 81 |  | 150 |

$$
\begin{array}{ll} 
& x^{2}=5.0648 \\
\text { a.f. }=9 & \quad .50>P>.30
\end{array}
$$

Where available data relate to order only or to a qualitative character capable of being ranked, the varjation of the scores between groups may be analyzed by method of ranks. By this method ranked data are used instead of the original values, avoiding the assumption of normality in the original data.

The method was suggested in a report on the validation of Gaston's A Test of 保icality. In that test the ovaluation of each pupil by the teacher did not result in a numerical score but in placing the pupil in one of $f$ ive categories. Ranks were assigned to the frequencies for each test score interval, and chi square rwas calculated. From this, the rank correlation ratio was calculated.

In the present study; the score intervals were the ratings. For each group of musicians pooled according to the anount of theory training, and for the criteition group, the ranks of the frequencies of cases In each score interval were determined, the largest frequency receiving the rank of 1. Since it had been found previously that the scores of the group with two years of theory and those of the group with three or more years of theory were homogeneous in distribution, the groups were combined for this study. The results appear in Table 20.

The sum of the sums of the ranke of the colums riad to be 135, that is $\frac{m n(n+1)}{2}$, where $m$ is the nuber of test groups and $n$ is the number of ratings (scores). If there had been complete lack of agreenent

[^32]
## TABLE 20

DISTRIBUTION OF TEST SCOLES ATM HANKS FOR MUSICIANS GROUPED
 CRITERION GROUP

| GROUP | Test Scores (Ranks in parentheses) |  |  |  |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | 6-19 | 20-29 | 30-39 | 40-49 | 50-59 | 60-69 | 70-79 | 80-89 | 90-102 |
| 1 yr. | $(21$ | (17) | $\begin{gathered} 21 \\ (3) \end{gathered}$ | (4) | ( 4 | ( ${ }^{2}$ | $(7.5)$ | $\begin{gathered} 0 \\ (9) \end{gathered}$ | $(7.5)$ |
| $2,3$ | (5) | (1) | (2) | $(35)$ | 19 (4) | $\begin{gathered} 6 \\ (7.5) \end{gathered}$ | $(7.5)$ | $\begin{gathered} 2 \\ (9) \end{gathered}$ | $\begin{gathered} 8 \\ (6) \end{gathered}$ |
| Crit. | $\begin{aligned} & 0 \\ & (8) \end{aligned}$ | $\begin{gathered} 0 \\ (8) \end{gathered}$ | $\begin{gathered} 0 \\ (8) \end{gathered}$ | $\frac{1}{(6)}$ | (3) | $\begin{gathered} 6 \\ (1) \end{gathered}$ | (5) | (3) | (3) |
| Rank totals | 15 | 10 | 13 | 13 | 12 | 14.5 | 20 | 21 | 16.5 |
| $(\text { R.T })^{2}$ |  | 100 | 169 | 169 | 143 | 210.25 |  | 4.7 | 272.25 |

between the ratings and the groups; the ranks would have been the samo for each rating and the sums would have been $3,6,9,12,15,18,21,24$, and 27, although not necessarily in that order. If there had been complete agreement, the sums of the colums would have been equal.

To determine whether or not there was complete agreement between the score ratings and the groups, the following formula was applieds

$$
X_{x}^{2}=\frac{12}{n p(p+1)} \sum(\text { Rank totals })^{2}-3 n(p+1)
$$

Where $n=3$ and $p=9$

$$
\begin{aligned}
& x_{2}^{2}=\frac{12}{3.9 \cdot 10}(2130.5)-3 \cdot 9 \cdot 10 \\
& x_{x}^{2}-93.742-90.000-3.742
\end{aligned}
$$

Entering the $X^{2}$ table yith 8 degrees of freedom, the probability Ievel for the value of 3.742 was found to 110 between .90 and .80 , which in this case was between .10 and. 20 , as the probabilities had to be read in the opposite direction. This finding indicated that there was substantial agreement between the scores and the groups.

The rank correlation ratio for the groups was estimated by the formula:

$$
\begin{aligned}
& n_{r}^{2}=\frac{x_{r}^{2}}{n(p-1)}=\frac{3.742}{3(9-1)}=\frac{3.742}{24}=.155916 \\
& r=1-\sqrt{.155916} \\
& r=1-.394=.606-.61
\end{aligned}
$$

Although the method of analysis was applied to data which necessitated the use of only three groups, the rank correlation of .61 which
was obtained appeared to Indicate that the variation of the score ranks wes essentially in accord with that expected for groups on fnoreasing levels of musical competence. Therefore the test was judged to possess validity, and was measuring that aspact of musicianship assoctated vith score reading.

## Additional Procedure for Validation

Having eatimated the validity of the test by one methody a further procedure was utilized in establishing its validity. Tests of the sigm nificance of the difference of the mean scores of the criterion group and each of the other groups vere caloulated. Each tatest was preceded by an F-test of homogencity of variance. Although a significant diference had already been found betmeen the mean of the group having one year of theory training and that having tro years, and no signsfioant difference was found between the mean of the group with two years of theory and that hayling three or more years, the comparisons between the criterion group and each of the other groups were nevertheless made in order to establish the Ievel of confidence at which the means could be said to be gignificantly different. The resulte obtained are shom in Table 21.

In each case, the assumption of homogencity of variance was met, since the obtained $T$ value for each comparison indicated that there were more than five chances in 100 that the disparyty between the calculated variances was due to chance. Hence the hypothesis was accepted that the samples came from the same population.

The obtained t ratio in oach comparison was significant, at the . 001 level. Indicating that there was less than one chance in 1,000 that

## TABLE 21

COMPARISONS EETVECN THE CRITERION GROUP AND EACH GROUP CLASSIFIED ACCORDING TO AFPUNT OF THEORY TRAINTNG

| Group | d.f. | F | P | Mean of Oriterion | Siean of Group | t | P |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 1 year of theory | 24,109 | 1.2871 | $>.05$ | 73.36 | 27.93 | 14. 1086 | $P<.001$ |
| 2 years of theory | 80,24 | 1.2494 | $>.05$ | 73.36 | 38.94 | 8.5814 | $p<.001$ |
| 3 or more yre. theory | 68,24 | 1.5473 | $>80$ | 73.36 | 44.34 | 6.5360 | $P<.001$ |

differences between the means as large as those obtained for the samples could have occurred by chance. Thus, the hypothesis that musicians rated as "above average" would achieve significantly better in the skill. measured by the teat than would unselected musicians was accepted, and the test was judged to be valid.

## Comparisons of School Harks in Music Courses

with Test Achievernent

The data for this part of the study were obtained from the subjects ${ }^{\text {( }}$ answers to Ho. 9 of the questionnaire. Thoy were to indtoate their achievement, according to most recent marks, in the theory courses spec ified. These wrere hamony, ear-training, bight singing, keyboard harm mony, or their combination into a general "theory" course.

The original intention in the mark comparisons was to categorize the subjects according to three classifications: (1) above average (marka of $A$ or $B$ or their equivalent), (2) average (a mark of $C$ or its equivam lent), and (3) below average (D or 7 or their equivalent). Howevery it was found that in all the courges there were too few marks in the "below average" category to obtain rellable results. Therefore, this category was combined with the "everage" category, and all comparisons then involved only two categories - - "above average" and "average or below".

Reliability of the Questionnalue Data on School Rearks

It was necessary first to detemine whether the course achievement marks indicated by the subjectes wore sufficiently accurate to rely upon. It was possible to check the actual marks of the subjects in only schools

A and $\mathrm{H}_{4}$ This rasy of course, not a randam smaple, but these sehools provided over onethird of the total test returns, and therefore the check on the reliobility of the responses mas believed to be sufficiently comprehensive for the purcose of this phase of the study. The courses for which mariss were available in both schools were hamony and ear-training. In school A marks were also available for the course in keyboard harmony.

Disregarding five tests for whioh school records provided no mariss it was found that for the most recent mark in harmony, 81 of the subjects checks agreed with the actual school record eark; and 21 diffored. Of these 21, seven had ohecked too high a grade categorys and fousteen too low. For the mott recent mark in ear-training, 85 of tho subjects ohecks agread with the sohool records, and 17 differed. of the seventeen whioh difeered, six had checised too hith a catagory, and aleven too lov. Thus, for the courber in harmony and ear-training, the total nomber of subjects' checks which egreed with the actual narise was 166, and those which aisagreed totaled 38. The proporition of accurate information by the subjects was, therefores approximately 77 percent. It was interesting to observe that approxinately twice as many subjects whose checks disagreed with the sohool racords undsrestimated their achievement as overestimited it.

A further aid in determining the accuracy of the information concerming achfevement in theory courses as reported by the subjects was to have been provided by the test administrators in tho various schools. In the special instructions for the test administrators a request mas made.
fom then to vertsy the eubjectst answers on quection we. 9, by enciroling the category which had beon, in thetr oplniou, wrongly checked, and then to checin the proper cetegory. Honever, onty a snall number of the retums from four of the schools had teen so checised.

For school A th pas posistble to verify the accuracy on such corractions by the test administrators. It wes found that of sixty corm rections made for harmony and ear-training marks, 41 afreed with the actual recorded marks, but 19 disagreed. Thats was only approxinately 50 per cent agreament, and therofora it was concluăd that the subjects remenbored better than the test adnfinstrators wint their most recent menhs had been, since approximetely 77 per cent of the subjecte had prom vicuan accurate information. Thereiong, eny corrections by aministrators were henceiorth disregarded.

A more necurate inverstigation of the reltability of the marks in theory courses reported by the subjects was made by corpartig the means of the scores faliing in eaoh mark category for each theory course, docording to reported achiovement and to actual aohievanent as found in school records. The resplts of the t-testes of the gimmificance of the Afference of the means are given In Table 22.

It was found for each comparison that any differences betreen the means arose only by chance. It qus therefore concluded that the achievement in theory courses according to mark oategories as roported by the subjects could be accepted. This was true, of course, only for schoola $A$ and $H$, but since these schools provided over onewthird of the total nuraber of test scores, there was little reason to belleve tinat the other returns would deviate from these findings to any appreciable extent.

COMPARISONS IN TEST ACHEVEMENT ACCORDING TO HARKS IN MUSIC THEORY COURSES AS REPORTED BY SUBJECTS AMD ACCORDTHE TO ACTOAL SCHOOL $H A R K S$

| hark Gategory | Theory Training status of subjects in years | Course | Schools | $\overline{\mathrm{X}}$ of Scores according to Warles Reported | $\bar{X}$ of Scores according to Actual Marks | d.f. | F | P | t | P |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| Above Average <br> Aver. or below | 1 | Harmony | A, H | $\begin{aligned} & 36.94 \\ & 32.19 \end{aligned}$ | $\begin{aligned} & 36.45 \\ & 31.40 \end{aligned}$ | $\begin{aligned} & 15,19 \\ & 15 ; 9 \end{aligned}$ | $\begin{aligned} & 1.12 \\ & 1.43 \end{aligned}$ | $>.05$ | $\begin{aligned} & .006 \\ & .031 \end{aligned}$ | $\begin{aligned} & >.9 \\ & >.9 \end{aligned}$ |
| Above Average Aver. or below | 1 | Ear Training | A, H | $\begin{aligned} & 43.08 \\ & 29.45 \end{aligned}$ | $\begin{aligned} & 43.53 \\ & 26.65 \end{aligned}$ | $\begin{aligned} & 11,14 \\ & 19,16 \end{aligned}$ | $\begin{aligned} & 1.25 \\ & 1.27 \end{aligned}$ | $>.05$ | $\begin{aligned} & .045 \\ & .658 \end{aligned}$ | $\begin{aligned} & >.9 \\ & >.4 \end{aligned}$ |
| Above Average <br> Aver. or below | 1 | Keyb. Harm. | A | $\begin{aligned} & 49.08 \\ & 26.00 \end{aligned}$ | $\begin{aligned} & 43.38 \\ & 26.09 \end{aligned}$ | $\begin{aligned} & 11,15 \\ & 14,10 \end{aligned}$ | $\begin{aligned} & 1.00 \\ & 1.62 \end{aligned}$ | $>.05$ | $\begin{aligned} & .601 \\ & .022 \end{aligned}$ | $\begin{aligned} & >.5 \\ & >.9 \end{aligned}$ |
| Ibove Average Aver. or below | 2 | Harmory | $\mathrm{A}_{\text {, }} \mathrm{H}$ | $\begin{aligned} & 43.80 \\ & 35.63 \end{aligned}$ | 43.10 <br> 37.17 | $\left\|\begin{array}{l} 29,29 \\ 17,38 \end{array}\right\|$ | $\begin{aligned} & 1.06 \\ & 1.59 \end{aligned}$ | $\begin{aligned} & >.05 \\ & >.05 \end{aligned}$ | $\begin{aligned} & .122 \\ & .367 \end{aligned}$ | $\begin{aligned} & >.9 \\ & >.7 \end{aligned}$ |
| Above Average Aver. or below | 2 | Ear Training | A. H | $\begin{aligned} & 48.10 \\ & 29.80 \end{aligned}$ | $\begin{aligned} & 46.87 \\ & 29.89 \end{aligned}$ | $\left\|\begin{array}{l} 28,30 \\ 19,17 \end{array}\right\|$ | $\begin{aligned} & 1.00 \\ & 1.09 \end{aligned}$ | $>.05$ | $\begin{aligned} & .224 \\ & .037 \end{aligned}$ | $\begin{aligned} & >.8 \\ & >.9 \end{aligned}$ |

TABLE 22 (continued)

| Mark Category | Theory Training status of subm jects in years | Course | Schools | $\overline{\mathrm{X}}$ of Scores according to Harks feported | $\overline{\mathrm{X}}$ of Scores according to Actual Yarks | d. 1. | F | P | $t$ | P |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| Above Average Aver. or below | 2 | Keyb. Harm. | A | 45.48 <br> 37.95 | 46.36 <br> 34.53 | $\begin{aligned} & 24,20 \\ & 21,18 \end{aligned}$ | $\begin{aligned} & 1.28 \\ & 1.38 \end{aligned}$ | $\left\|\begin{array}{l} >.05 \\ >.05 \end{array}\right\|$ | $\begin{aligned} & .135 \\ & .915 \end{aligned}$ | $\begin{aligned} & >.8 \\ & >.3 \end{aligned}$ |
| Above Average <br> Aver or below | $3(+)$ | Hamony | A, H1 | 39.70 <br> 39.64 | 44.64 $29.71$ | $\begin{aligned} & 13,9 \\ & 10,6 \end{aligned}$ | $\begin{aligned} & 2.31 \\ & 4.90 \end{aligned}$ | $\begin{aligned} & >.05 \\ & >.01 \end{aligned}$ | $\begin{aligned} & .639 \\ & 1.010 \end{aligned}$ | $\begin{aligned} & >.5 \\ & >.3 \end{aligned}$ |
| Above Average <br> Aver. or below | $3(+)$ | Ear Training | A, H | 48.73 $30.00$ | $\begin{aligned} & 47.62 \\ & 26.17 \end{aligned}$ | $\begin{gathered} 12,10 \\ 7,5 \end{gathered}$ | $\begin{aligned} & 1.01 \\ & 4.06 \end{aligned}$ | $\begin{aligned} & >.05 \\ & >.05 \end{aligned}$ | . 112 .634 | $\begin{aligned} & >.8 \\ & >.5 \end{aligned}$ |
| Above Average Aver or below | $3(+)$ | Keyb, Hamm. | a | $\begin{aligned} & 49.140 \\ & 29.00 \end{aligned}$ | $\begin{aligned} & 44.14 \\ & 27.67 \end{aligned}$ | $\begin{aligned} & 13.9 \\ & 9.5 \end{aligned}$ | $\begin{aligned} & 1.12 \\ & 1.87 \end{aligned}$ | $\begin{aligned} & >.05 \\ & >.05 \end{aligned}$ | $\begin{aligned} & .610 \\ & .218 \end{aligned}$ | $>.5$ |

Comparisons in Test Achievement between Bubjects Receiving "Above Average" and Subjects Receiving "Average or Below"

## Marks in Music Courses

Achievement, according to the nost recent mails, in the music theory courses specified in the questionnaire was taken to be that indicated by the subjects, with the exception of the subjects in schools A and $H_{s}$ for which the actual marks were available. The tests having no data provided on mariss in music courses were onitted. The schools which provided data for the study are indicated in Table 23. No study of the relationship of marks in sight singing to test achievement was possible, since this course was not offered as a separate course in the schools Included in the study. Only two schools offered ifist year keyboard harmony, and only two offered first year "theory". Only school a offered second year keyboard haxwony. There were insufficient data to justify a study for those schools offerthe second year "theory".

T-tests of the significance of the differences between the means of the scores of the subjects whose marke were "above average" and those whose marks nere "average or below were calculated for all the theory courses for which there were sufficiont data. The comparisons and results are given in Table 23. F-tests for homogeneity of variance prem ceded each t-test. In all but three of the tests the calculated ratios of the variances indicated that the assumption of homogenaity vas not met, and the d-test was applied.

For the subjecte heving one year of training, there was a significant difference in test achievement between those recaiving above average

COMPARISONS IN TEST ACIIEVEMETP BETAEEN SUBJECTS REGETVING＂ABOTE AVERAGE，$n$ AMD SUBJECTS RECEIVING＂AUERAGE OR BELOM＂WARKS IN EUSIC THEORI COURSES

| Tears of theory training | Course | Schools <br> Fooled | d．f． | F | $p$ | $\bar{X}$ of ＂AA＂ | $\bar{X}$ of HAOB | $t$ | Prabability |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 1 | Hammay | A， $\mathrm{O}, \mathrm{E}, \mathrm{H}$ | 29,34 | 3.34 | $<.01$ | 34.37 | 25.11 | $1.975 *$ | ． $05>P \gg .01$ |
|  | Ear Training | A，可层， H | 28，35 | 8.63 | $<.01$ | 38.07 | 22.39 | 3．270\％ | $<.01$ |
|  | Keyb．Harmony | 1， $\mathrm{D}^{\text {d }}$ | 21，21 | 8.68 | $<.01$ | 41.05 | 23.77 | 2．945\％ | $<.01$ |
|  | ＂Theory＂ | B，G | 21，11 | 1.27 | $>.05$ | 32.75 | 29.18 | 3.430 | $<.01$ |
| 2 | Harmomy | A，B，D，H | 37，36 | 2.682 | $<.01$ | 42.71 | 35.05 | 1．616\％ | $>.05$ |
|  | Ear Training | A，B，D，H | 35，35 | 4.474 | $<.10$ | 46.56 | 30.58 | 2．115゙＊ | $<.01$ |
|  | Keyb．Harmony | A | 27，28 | 4.23 | $<.01$ | 46.89 | 34.68 | 1．694\％ | ． $05>P \gg .01$ |
| 3 or more | Harmony | A， $\mathrm{C}, \mathrm{F}, \mathrm{O}, \mathrm{H}$ | 29，27 | 2.05 | $>.05$ | 47.67 | 41.29 | 1.080 | $>.05$ |
|  | Ear Training | A，C，F，H | 28，26 | 3.11 | $<.01$ | 55.24 | 34.04 | 1．701＊ | $<.01$ |
|  | Keyb．Harim． | A | 13，5 | 4.16 | $>.05$ | $W_{4} .74$ | 27.67 | 1.785 | $>.05$ |

＊Behrens－Fisher d－test．
marks and those receiving average or bolow average marks in the courses in ear-training, keyboard harmony, and theory as an integrated course. The difference obtained between the high and Iov ranking groups in harmony was in the region of doubty, i.e., between the . 05 and . 01 levels of confidence.

For the subjects having twa years of training, there was a sigw nificant difference in test achlevement between high and low ranking groups in the course in ear-training. The finding for keyboard harmony was in the region of doubt. There was no gignificant dfference in test achievement according to marks in harmony.

For subjects having three or more years of trainings there was a significant difference in test achievement between high and low ratod groups only in ear-traininge. In harmony and keyboard harmony no significant difference was found. These subjects were, of course, one or nore years removed from the time of their taking these courses, but even so the marks in eer-training at that time apparently predicted their achievement on the tost.

For any leval of anount of training, then, the mariss in the eartraining course correlated most closely with achievenent on the test. Ear-training is very closely related to the skill demanded by the test, that $4 s$, reading and hearing intervals and chords. Since achievement in ear-training courses according to school marks corresponded significantly with test achievement, this provided additional proof of the validity of the test.

Marks in hamony, except the doubtful finding for the subjects heving one year of training, did not correspond significantly with
achievement on the teat. This was to be oxpected, since marks in hamony are often assigned on the basis of achievement in activities not related to a functional use of knomleage of hamony. The finding substantiated the crithcism of harmony courses in mich emphasis is placed, for instance, on the mechanical mriting of hamonies divorced from actual musLeal values or other practical considerations.

The signiffeant relationship found botween test achiovement and marks in keyboard hamomy anong the subjects heving one year of training, and possibly for those having two years, incicated that lnowledge of the keyboard and facility in playing chorda mas important.

Piano players are accustomed to reading both treble and bass staves simultaneously, and they deal mith intervals and ohords both by sight and muscular feel. The relationship of experience on the keyboard and of leyboard facility to achervement on the test were investigated later in the study.

> The Relationship of Selected Factors to
> Achievement on the Test

In the iollowing comparisons the data for all subjects yere pooled. In each comprison, onty the scores of the subjects providing. date pertaining to the characteristic were used.

Chi square was used to test the hypothesis that any observed differences between groups in the distribution of scores were due to chance, i.t., that the factor under consideration had no slgnificant relationship to test achievement.

The tabulation of scores acoording to age yielded large groups for the ages of $18,29,20$ and 21. Below the age of 18 and above the age of 21 there were too few scores to deal with separately. Therefore, those of age 17 were included with age 28 , and from age 22 on all were erouped together. The results of the test of homogeneity of distribution are given in Table 2he

The obtained chi square value proved to be too small to indioate a significant difference in the distribution of scores, and the hypothesis that there was no difference in test achievement according to age was accopted. A trend was indeated, since the probability lay between .20 and .10, buth an examination of the expected cell frequencies traticated that cell $c$, with too many observed frequencies, contributed a disproportionate anount to the ehi square total. Oell j, with too many observed frequencies, and cell $u$, with too few, also contributed excessively to the total. If there had been a true trend such anounts would more likely have occurred in the cells of the "above 21" age column.

## Sex

The distribution of ecores according to sex, and the results of the chi square test are show in Table 25. The obtained ohi square value indfcated that the differences in the distribution were due to chance, and the hypothesis of no difference mas confidently accopted. The factor of sex apparently had nothing to do with the test results.

## TABD: 24

ANALYSIS OF HORCOEMETY OF DESTRTBETION of scoris accordma to acb

| Scores | 17-18 | 19 | $\begin{aligned} & \text { Ages } \\ & 20 \end{aligned}$ | 21. | above 21 | Totals |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| $50 \%$ | 6 (a) | II (b) | 16 (c) | 7 (d) | 7 (e) | 47 |
| 40-49 | 6 (f) | 7 (5) | 11. 3 ) | 4 (k) | 6 (1) | 34 |
| 20-39 | 30 (a) | 43 (n) | 19 (0) | 20 (p) | 23 (q) | 135 |
| -19 | 12 ( r ) | 14 (s) | 5 (t) | $4(u)$ | $4(\mathrm{v})$ | 39 |
| Totals | 54 | 76 | 51 | 35 | 40 | 255 |

$$
\text { d.f. } 12 \begin{array}{ll}
x^{2}=18.036 & 20>P>.10
\end{array}
$$

## TABLE 25

ANALYSIS OF HOROGEMEITI OF DISTRIBUTION OF SCORES ACCORDING TO SEX

| Scores | Women | Men | Totals |
| :---: | :---: | :---: | :---: |
| 40 and above | 56 | 26 | 82 |
| $20 m 39$ | 91 | 48 | 139 |
| 19 and below | 22 | 17 | 39 |
| Totals | 169 | 91 | 260 |

$$
\begin{array}{cc}
x^{2}-1.669 \\
\text { d.f. }=2 & .50>P>.30
\end{array}
$$

## Voice Part Usually sing in Ensemble

A factor which could have affected the test results was whether a person sang soprano, alto, tenor or bass. A person might tend to concentrate on the part which he is accustomed to reading and singing in onsemble. In the test, in order to introduce errors in performance which were not obvious, there were nore errors of performance in the inner and bass parts than in the soprano part. Fossibly an advantage could accrue to singers on those parts. If there was a difference, then the test would not consist of homogeneous items, and many of the assumptions on which the reliability and validity of the test were based would have had to be abandoned. Therefore, the scores were classified according to the part sung by the musicians, and were subjected to the test of homogeneity of distribution. The results of the analysis given in Table 26 indicated that any observed differences in the distribution of scores arose only by chance, and that therefore the part sung had no relation to the scores which were obtained on the test.

Principal Medium of Performance

In order to determine whether there was any difference in achievement on the test according to the principal medium of performance, the scores were tabulated for the principal media of piano, voice, and instrument (other than piano). The results of the analysis of the distribution of scores are given in Table 27.

The probability that a chi square value as large as that obtained

## TABLE 25

 AcCORDIN TO THE PART SUHG

| Scores | Soprano | Fart <br> Tenor | Bass | Totals |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
| 404 | 26 | 30 | 10 | 16 | 82 |
| $20-39$ | 45 | 39 | 20 | 27 | 131 |
| -19 | 14 | 8 | 6 | 10 | 38 |
| Totals | 85 | 77 | 36 | 53 | 251 |

$$
x^{2}-6.438
$$

$$
\text { d.f. }=6
$$

$$
.50>P>.30
$$

## 




| Score | Plano | Volca | Instrunert <br> (String or wind) | Totals |
| :---: | :---: | :---: | :---: | :---: |
| $40+$ | $47(\mathrm{a})$ | $17(\mathrm{~b})$ | $7(\mathrm{c})$ | 71 |
| $20-39$ | $43(\mathrm{~d})$ | $41(\mathrm{o})$ | $49(\mathrm{~s})$ | 133 |
| -19 | $11(\mathrm{~g})$ | $12(\mathrm{~h})$ | $33(\mathrm{i})$ | 56 |
| Totale | 101 | 70 | 89 | 260 |

$$
x^{2}=4.494
$$

$$
\text { d.f. }=4
$$

$P<.001$
could occur by chance was Iess than I in 1,000 , and the hypothesis that the distribution was homogeneous was rejocted. Examination of individual cell contributions to the chi square total revealed that cell a contributed far too much, and cell g too little; cell c had too fev observed frequencies and cell 1 too many. Tmus it appearec that the musicians Whose principal instrument was piano achieved significantly better scores on the test than did the instrumentalists, and possibly also better than the vocalists.

This finding was investigated further by comparing the distribum tions of the scores of the piano majors with those of the voice majors, and with those of voice and instrumental majors combined. The results are given in Tables 28 and 29.

The probability of a difference in test achievenent between pianists and vocalists lay betreen the . 02 and . 01 levels of confidence. This was regarded as sufficiontly significant to indicate that the musicians who concentrated particularly on the piano did better on the test then the vocalists because, normally, singers need to use the piano to some extent. Xany voice majors even have the pieno as a second major performing medium. Without this factor of considerable piano experience, then, the voice majors would probably have achieved significantly worse than the piano majors on the test.

When the piano majors were compared with all others combined, a significant difference in the distribution of scores was found at the . OOI level of confidence. Musicians whose major medium of performance was piano apparently achieved significantly better on the test.

## TABLE 28

## ANALYSIS OF HONOGENEITY OF DISTRIBUTION OF SCORES OF PLATO RAJORS AND VOICE MAJORS

| Score | Plano | Voice | Totala |
| :---: | :---: | :---: | :---: |
| $40+$ | 47 | 17 | 64 |
| $20-39$ | 43 | 41 | 84 |
| -19 | 11 | 12 | 23 |
| Totals | 101 | 70 | 171 |

$$
\begin{array}{ll}
x^{2}=8.824 \\
d . f . & \\
& .02>P>.01
\end{array}
$$

## TABE 29

##  <br> OF PTANO MAJOES ARD AJL OTHESS

| Scores | Manists | All otrers | Totals |
| :---: | :---: | :---: | :---: |
| $40+$ | 47 | 24 | 71 |
| $20-39$ | 43 | 90 | 133 |
| -19 | 11 | 45 | 56 |
| Totals | 101 |  | 260 |

$$
x^{2}-33.427
$$

$$
\text { C.f. }=2 \quad \mathrm{P}<.001
$$

## Piano Experience

AIl subjects, disregarding principal perforning medium, were classtfied according to the number of years of piano playing experience, ranging from none to eight years and above. When the number of years of experience was not provided in question Ho. 6 of the questionnaire, the years of study mere taken, or the reverse. The obtained data were best condensed into four categoriest one year or less of experience, two and three years, four and five years, and six or more years. The results of the comparison of the distribution of scores are given in Table 30 .

Since the probebility was less than one in 1,000 that a ohi square Value as large as that obtained could have occurred by chance, the hypothesis of homogeneity of distribution was rejected. The data appeared to indicate that scores improved as the number of years of experience on the piano increased. Analysig of cell frequencies and their contribution to the obtained chl square value indicated that this was true particularly after slx and more years of experience.

The finding was investigated further by comparing the score aistributions of the subjects in the first three categories of piano exper lence. The results are given in Table 31.

Although a trend possibly was indicated, the kypothesis of homom geneity was accepted, since the probability of observed differences occurring by chance was between. 20 and . 10 .

It was concluded that scores improved gignificantly after approximately six years of plano experience, but up to approximately six years there was no signifioant inprovement in achievement on the test. The

## TABLE 30

AMALYSIS OF HOUCGEUEITY OF DISTRIBUTION OF SCORES AOCODDIMG TO MOBBER OF YEARS OF PIAHO PLAYTHG EXPERIENCE

| Score | $0-1$ | Years of experience <br> 2-3 $4-5 \quad 6$ and more |  |  | Totals |
| :---: | :---: | :---: | :---: | :---: | :---: |
| $40+$ | 7 | 4 | 9 | 64 | 84 |
| 20-39 | 15 | 30 | 19 | 66 | 130 |
| -19 | 10 | 7 | 9 | 13 | 39 |
| Totals | 32 | 41 | 37 | 143 | 253 |

$x^{2}=29.216$

$$
d_{0} f_{-}=6 \quad P<, 001
$$

## 

AMALTSIS OF HOHOGEIETIS OF DISTETBUTION OH SCORWS OE SUBJECTS HAVME TP TO FTVE YEABS OF FIANO ELAYTNG EXPERTMNE

| Scores | Tears of experience |  |  | Totals |
| :---: | :---: | :---: | :---: | :---: |
| $40+$ | 7 | 4 | 9 | 20 |
| 20-39 | 15 | 30 | 19 | 64 |
| -19 | 10 | 7 | 9 | 26 |
| Totals | 32 | 41. | 37 | 110 |

$$
x^{2}=6.731
$$

$$
\text { d.f. }-4 \quad .20>P>.10
$$

contribution of keyboard experience to the skill measured by the test apparently was slight for several years, but became much greater after five or six years.

## Piano Study

The subjects indicating that they had had some amount of formal piano study were classified into the following categories: one year or less, two and three years, fout and five years, and six or more years of study. If only piano experience was indieated, nad no study; the scores of those subjects were net included in tine comparisono.

Since the analysts of the effect of the amount of piano experience had indicated that approximately six or nore years of experience made a significant difference in the distribution of scores, it was decided to compare the distributions in the fitrst three above categories first, and then to compare their pooled distribution with that of the last category. The results are given in Tables 32 and 33.

As found for piano playing experience, the results indioated that up to five years of fomal study made no significant difference in the distribution of scores, although a trend was possibly inaicated. Six or more years of study resulted in a vomy significant difference in achieve ment on the teet.

## Voice Study

The distribution of scores accoraing to the anount of Cormal voice study was tabulated only for voice majors. These were pooled into the following categories: one year or less, two and three years, four and

TABLSE 32

ANALYSIS OF HOMOGMEITY OF DISTRIBUTION OM SCORES OF SUBJECTS HAVING UP TO ETVE YEARS OF PLANO STUDY

| Scores | 1 or IessYears of Study <br> $2-3$ | $4-5$ | Totals |  |
| :---: | :---: | :---: | :---: | :---: |
| 404 | 9 | 8 | 11 | 28 |
| $20-39$ | 17 | 39 | 19 | 75 |
| -19 | 11 | 9 | 6 | 26 |
| Totals | 37 | 56 | 36 | 229 |

$$
x^{2}=7.598
$$

$$
\text { d.I. }-4
$$

$$
.20>P>.10
$$

## TABLE 33

COMPAMISON B HTHEN THE DTSTREBUTION OF POOLED SCORES OF SUBJECTS KAVIMG UP TO FIVE TEARS AM THAT OF SUBJECIS HAVING SIX OR RJORE YBARS OF PIANO STUDY

| Scores | $\begin{aligned} & \text { Years of Study } 6 \text { or more } \\ & 5 \text { or less } \end{aligned}$ |  | Totals |
| :---: | :---: | :---: | :---: |
| $40+$ | 28 | 55 | 83 |
| 20-39 | 75 | 52 | 127 |
| -19 | 26 | 11. | 37 |
| Totals | 129 | 118 | 247 |

$$
x^{2}=18.576
$$

$$
\mathrm{a}_{.} \mathrm{f} \cdot 2 \quad \mathrm{P}<.001
$$

five years, and six or more years of study, The distributions and rem sults of the test for homogeneity are shom in Table 3 .

The obtained chi square value indicated that the distributions were homogeneous, any observed differences occurring only by chance. It was therefore concluded that the skill measured by the test did not improve through continued volce stuog. This finding was to be expected, since volce study is primarily concemed with problems of vocal production. Also, the single melodic line is stressed and not the effect of vertical tonal relationships.

## Voice mperience

No. 7 of the questionnaire, as for the preceding study provided the data to place the subjects in categories of years of singing experience. Again, only the test results of voice majors were considered. The categories were condensed so that more scores could be obtained for each. Wherever the anount of singing experience was not indicated in the questionnatre, the number of years of study was taken to indicate the anount of experience. The comparison of the score distributions is given in Table 35.

No difference, other than chance, was Lound in the distribution of scones, and the hypothesis of homogeneity was accepted. Apparently the number of years of singing experience of voice majors played no significant role in the skill measured by the test.

Piano Experience of Voice Majorg: - . Since, for all subjects pooled, piano experience and study had proved to have a significant

## TABLE 34

ANLIESTS OF HOMOGMBEITY OF DISTRIBUTION OF SCORES OF VOICE WAJORS Acconding to Mubre of zeans or Voree sivil

| Scores | 1 or less | $\begin{array}{r} \text { Tears } \\ 2-3 \end{array}$ | $\underset{4-5}{4-5 d y}$ | 6 or mare | Totals |
| :---: | :---: | :---: | :---: | :---: | :---: |
| 40* | 2 | 9 | 5 | 1 | 17 |
| 20-39 | 9 | 11 | 6 | 15 | 41 |
| -19 | 1 | 4 | 3 | 4 | 12 |
| Totals | 12 | 24 | 14 | 20 | 70 |

$$
\begin{array}{cl}
x^{2}=8.053 \\
\text { a.f. }-6 & \quad .30>p>20
\end{array}
$$

## TABIE 35

AKALISIS OF HOLOOBNEITY OF DISTRIBITION OF SCORES OP VOICE MAJORS ACCORDING TO TEARS OF SINGIMG EXPERIBNCE

| Scores | 7-5 | Years of mxperience $6-7 \quad 8$ or more |  | Totals |
| :---: | :---: | :---: | :---: | :---: |
| $40+$ | 6 | 9 | 6 | 21 |
| 20-39 | 12 | 8 | 17 | 37 |
| -19 | 5 | 4 | 3 | 12 |
| Totals | 23 | 21. | 26 | 70 |

## $x^{2}-4.134$

$$
\text { d.s. } 4
$$

$.50>P>.40$

Influence upon the skill measurer by the test after a period of approxImately six yeare, it was decided to detemnine whether that held true for voice majors nione. The roice majors were classirted into categories eccording to years of piano stuay or experience, whichever was greater. These oategories of experience were condensed as shom in Table 36, and the hoinogeneity of the dintributions of the scores wes tested.

Since the result vas that there were more than ten enawces In 100 that any obserged dfferences vere due to chance, the hypothesis of no difference was accepted, although possibly a trend was indicated.

However, the hypothesis could be rejected to the .05 level of confidence when the subjects were pleced in two broad categortes of six years or less, and seven years or more, of piano experience, as blom in Table 37. This closely parslieled the similar finding of the offect of piano exparience on test results for all subjects.

Principal Instrument, Other than Piano, of Instrumental Majors

This phase of the strdy wes to determine whother there was any difference in test achievement between the mustcians whose principal instrument was string, bress, or woodind. Subdivisions of the groups into specific instrumente, such as violin or trumet, etc., were not feasible because of the scattered data thus obtained. Only one subject was a peraussion major, so a percussion caterory was not possible. The distributions of scores in the three groups, and the resuits of the anolysts are given in Table 38.

## TABLE 36

AJaLISIS OF HOHOGMEITY OF DISTRIBUTION OF SCORES OF VOICE MAJORS ACCORDIN TO ALOOUNT OF PIAMO EXPBRIENCE

| Scores | Years of Piano Exparience4 or less4 or more |  |  | Totals |
| :---: | :---: | :---: | :---: | :---: |
| $40+$ | 4 | 2 | 12 | 17 |
| 20-39 | 15 | 13 | 13 | 42 |
| -19 | 5 | 4 | 3 | 12 |
| Totals | 24 | 19 | 27 | 70 |

$$
x^{2}=6.709
$$

$$
\text { d.f. }-4
$$

$$
.20>P>.10
$$

## TABLS 37

## ANALYSIS OF HOMOCREITY OF DISTRIBUTION GF SCORES OF VOICE

 HAJORS ACCORDIIG TO BROAD GATEGORIES OF PIANO EXPERTBHCE| Scores | Tears of Plano Experience <br> 7 <br> or less or more | Totals |  |
| :---: | :---: | :---: | :---: |
| 407 | 6 | 11 | 17 |
| $20-39$ | 28 | 13 | 41 |
| -19 | 9 | 3 | 12 |
| Totals | 43 | 27 | 70 |

$$
\begin{aligned}
& x^{2}=6.6119 \\
& \text { d.f. }=2
\end{aligned} \quad .05>\mathrm{P}>.01 .
$$

## TABLE 38

AMALISIS OF HOLOGENISTTI OF DISTRIBUTION OF SCORES OF STAING, BRASS, AND WOODITMO WHJORS

| Scores | String Principal Instriment | Brass | Woodrind | Totals |
| :---: | :---: | :---: | :---: | :---: |
| 404 | 3 | 7 | 9 | 19 |
| $20-39$ | 12 | 22 | 19 | 53 |
| -19 | 3 | 10 | 4 | 17 |
| Totals | 18 | 39 | 32 | 89 |

$$
x^{2}-2.962
$$

d.f. $=4$
$.70>P>.50$

The distributions were found to be homogeneous, and it was therefore concluded that there was no difference between string, brass, and woodwind players in their achievement on the test.

Since the scores were found to be homogeneous in distributions, the subjects wore pooled in the following comparisons.

Amount of Study of Erincipal Instmment by Instrumental kajors

The pooled instrunental majors were classified according to the length of time of etudy on their principal instrument. Table 39 gives the results of the analysis of homogeneity of the distribution of scores.

The groups were found to be similar in test achievenent, any observed differences in the distribution of their scores occurring only by chance. As with the finding for the effect of the anount of study of voice by voice majors, this was interpreted to mean that study on instruments which are essentially melodic does not contribute to the skill measured by the test.

## Length of Experience on Wajor Instument <br> of Instrumental Majors

The pooled instrumental majors were classified into two broad categories of experience on their principal instruments, as follows: seven or leas, and eight or more years of experience. These broadly inclusive categories were used in order to obtain sufficient data in each category. The distributions of scores and results of the analysis are given in Table 40.

## TABLE 39

ATALYSIS OF HOEOMBETHY OF DISTRTBUTION OF SCORES OF INSTRUMENLAL HAJORS ACCORDHH TO AYOUNT OF STUDY ON THETE PRTNOLPAL TMSTRUUENE

| Scores | 3 or Iess | Years of 4 to | 7 or more | Totals |
| :---: | :---: | :---: | :---: | :---: |
| 404 | 5 | 5 | 9 | 19 |
| 20-39 | 16 | 16 | 21 | 53 |
| -19 | 10 | 3 | 4 | 17 |
| Totals | 31 | 24 | 34 | 89 |

$x^{2}-5.692$

$$
d . f . \in 4
$$

$.30>\mathrm{P}>.20$

## TAELE 40

ANALTSTS OF HOROQENETH OF EISTRTBUTON OF SCOHE OF TMETRU-



| Scores | Years of Expertence7 or lass mione |  | Totals |
| :---: | :---: | :---: | :---: |
| 404 | 5 | $\mathrm{I}_{4}$ | 19 |
| 20-39 | 22 | 31. | 53 |
| -19 | 11 | 6 | 17 |
| Totals | 38 | 51 | 89 |

$x^{2} * 5.480$

$$
\text { d.s. } 2
$$

$p>.05$

The distributions were fcund to be homogeneousg and therefore the concluston was the same as that for the amoun of stady of the principal instrument, nanezy, that increased exiexience on ratodic inetruments dees not contribute significantly to the sinil measured by the test.

Arount of Piano Expergence of Instrumente Maforst - Inasmuch as piano experience hac boen found to have e sigmificant relabionship to test achtevenent for all subgects when pooled, perticulariy after six or more years, and an amost sloiler mesult had been found for voice majors, the offoct of piano experience was studied tith the instrumental tajors alone. They were classif?ed into categories of anounts of piano expexience, and the distribution of the scores was analyzed. The results are given in Table 41.

It was found that any differences in the distributions of scores of the categories were due only to chance, and the hypothesis of homogeneity of the distributions was accepted. For instrumental majors, the anount of plano experience did not appear to make any slgnificant difference in test achievement. Although this finding was contrary to that found for all subjects pooled, and for voice majors, this may be due to the fact that the instmmental majors, according to the questionnaire data, tended to play, as secondary instruments, other non-keyboard instruments much more than they did the plano. Possibly they did not consider the piano to be as useful to them, or they were not as interested in playine the piano. There is also the possibility that, followIng the curriculum for non-keyboard instrumental majors, they were rem quired to study other instruments, leaving comparatively little tine for experience on the piano.
qABLE 41

ANALYBIS OF HOMOCBNETIT OF DISTRIBUTON OF POOLED SCORES OF IMSTRUNEWLAL MAJORS CLASSIFTED AGCORDING TO A OF PTANO EXPETIENGE

| Scores | Tears of Plano Experience <br> 1 or less $\quad 2-3 \quad 4-5 \quad 6$ of more |  |  |  | Totals |
| :---: | :---: | :---: | :---: | :---: | :---: |
| $40+$ | 6 | 2 | 5 | 6 | 19 |
| 20-39 | 11 | 35 | 8 | 15 | 49 |
| -19 | 8 | 4 | 2 | 3 | 17 |
| Totals | 25 | 21. | 15 | 24 | 85 |

$$
x^{2}=6.958
$$

$$
d_{.} f_{.}=6
$$

$$
.50>\mathrm{P}>.30
$$

The effect of piano sight reading ability upon test achievenent was estimated by the simple standard of thether the subject could or could not play hym tunes and simple accompaninents at sight. The data for this study was provided in No. 6 of the questionnaire. The subjects were placed in these two categories, and the diatributions of the scores were compared. The resulta are givon in Table li2.

A very significant difference was found between the scores of thase able to play simple material at sight on the piano and the scores of those not able to do so. Abjlity at the keyboard in tems of playing at sight the notation of simple materlal, similar to much of the naterial contained. In the test; appeared to be definitely assooiated with higher tegt achievement.

Piano Tmproviging Ability

The relationship of piano improvising ability to achievenent on the test was estinated by comparing the scores of the subjects who could, and those who could not, inprovise some type of accompaniment for sinaple melodies. The results are shom in Table 43.

The skill necessary to inprovise an accompaniment presunably requires some knowledge of the sound of at least the basic chords. But whether a reading knowledge of their notation is necessary is open to question, gince $^{\text {mary }}$ persons can improvise quite well mithout being able to read notation at 211. However, skill in reading at least the notation of the melody would seon to be necessary if an accompaniment is to

## TABLE 42

COMPARISON OF TEST ACHILVEUENT OF SUBEECTS ABLE, AND SUBJECTS NOT ABLE, TO PLAE STMPLE PIANO WUSTC AT STCHT

| Scores | Able | Not able | Totals |
| :---: | :---: | :---: | :---: |
| $40 *$ | 78 | 6 | 84 |
| $20-39$ | 107 | 24 | 131 |
| -19 | 25 | 14 | 39 |
| Totals | 210 | 44 | 254 |

$$
x^{2}-15.654
$$

d.I. $=2$
$P<.001$

## TABLE 43

COLPARISONS BETUEEN TEST SCORES OF SUBJECTS ABTE, ATD UMABIE TO IUPROVISE PIANO ACCOMPANDEENS FOR SIMPLE EPLODIES

| Scores | Able | Unable | Totals |
| :---: | :---: | :---: | :---: |
| $40+$ | 69 | 12 | 61 |
| $20-39$ | 98 | 30 | 128 |
| -19 | 21 | 13 | 34 |
| Totals | 188 | 55 | 243 |

$$
x^{2}=7.145
$$

$$
\text { d.f. }-2 \quad .05>P>.02
$$

be provided. In any event, all the subjects parkieipating in this studys, being masic majors, could presumably read music notation. Many of them had had some training in keyboard hermonys in which course a functional application of knowledge of chord structure and chord relationships to hamonizing nelodies is practiced.

The result of the comparison of the two groups was in the region of doubt, $1 . e$. , between five and two chances in 100 that the observed differences between the scores of the groups were due to chance. This strong trend toward a statistically significant difference was interm pretad to mean that inprovising skili for musicians who all tead music had more than just a chance relationship to the silll measured by the test.

## Choral Directing Experience

Only fifty-three subjects to whon the test was administered had had any choral directing experience. Of these, thirty-one had one year or less of experience, and the remainder had varying amounts ranging up to a number of years. To deal with these data, the subjects were placed in only two categories: one year or less of directing experience, and two or more years. The results of the comparieon of the score distributions of the groupe are given in Table W.

The distribution of scores was found to be homogeneous. 萑ith the data available, an increase in the number of years of choral directing experience did not result in a significent improvement in the skill measured by the test. Additional test data might possibly modify this

## TABIE 44

ANALYSTS OF HOMCEBMETT OF DISTRIBUTION OF SCORES ACCORDHG


| Scores | Years of Experience <br> I or less | or more | Totals |
| :---: | :---: | :---: | :---: |
| $40+$ | 11 | 11 | 22 |
| $20-39$ | 18 | 9 | 27 |
| -19 | 2 | 2 | 4 |
| Totals | 31 | 22 | 53 |

$$
x^{2}=1.53
$$

$$
\text { d.f. } 2
$$

$.50>P>.30$
finding, but it is recalled that an increase in ase, with correspondingly more general experience with maic, also did not result in a significant improvement in test achievement.

Instrumental Directing Experience

Only thirty-four subjects indicated that they had had any instrumental directing experience, and of these seventeen had only one year or 1ess. The romainder had varying amounts. As with the preceding study, the subjects were placed in tro categories, one year or less of experience, and two or nore years of experience. The results of the analysis of the score distributions are given in Table 45.

The hypothesis of homogeneous distributions of scores was again accepted. An increase in the number of years of instrumental directing experlence, according to the avallable data, did not result in increased achievement with the skill moasured by the test.

The Method Used Predominantly in Histeming
for Accuracy of Hamonic Effect

No. 10, part a, of the questionnaire, which was answered by the subjects after the administration of the test, was to provide data to determine which of geveral obvious methods of listening was used predominantiy in determining the accuracy of harmonic effect. The mothods were: (1) listening to one or two parts at a time (melodic analysis), (2) listening to the complete chord quality, such as a chord being major, minor, etc. (harmonic analysis), (3) combination of methods one and two, (4) listening primarily for appropriateness of harmonic effect and style,

## TABLE 45

ANALYSIS OF HOROGZNEITY OF DISTRIBUTION OF SCORES ACCORDTNG TO YEARS OF TMSTRUWMTAL DIRECTING EXPERTENCE

| Scores | Tears of Experience <br> 1 or less | Tor more | Totals |
| :---: | :---: | :---: | :---: |
| $40+$ | 5 | 7 | 12 |
| $20-39$ | 11 | 8 | 19 |
| -19 | 1 | 2 | 3 |
| Totals | 17 | 17 | 34 |

$x^{2}=1.140$
d. $\mathrm{I}_{0}=2$
$.70>P>.50$
thin melodic flow (no techaicel analigsis).
The intent of the study vac to edeterning generally whethor mothods 1, 2 or 4 trare used the mosts hat it was unticlpated that meng subjects would use both method 1 and method 2 to some extent, if thay used then at all, and that they minght be wable to decide which of the tro methods they used the more. Thexefore method 3 was included; but if they checked thes, they were still to indicate in the sub-sections provided for mothod 3 whether method 1 or method 2 was predominant in the cembination.

In tabulating the data, all ambiguous data were discarced. Method 3 was the one most often checked by the subjects, but these data were not used uniess the subjects had indicated in the oub-sections for that method that method 1 or 2 had been used predominantly. Therefore, the method indicated as boing used was in most fnstances only the predominant ore, and not that method exclusively. The anelysis of the obtained data is given in Table 46.

The hypothesis of a homogeneous distribution of test scores was rejected at the . 01 level of confidence. Apparently the predominant use of one of the methods made a sigrificant difference in test achievement. Analysis of the Andividual celle showed that cell $\mathrm{b}_{\text {, }}$ with too many obseaved frequencies, and cell $h$, with too fev, contributed excessively to the obtained chis square value. Both of these occurred under method 2 , analysis of complete chord qualifty.

To verify the finding further, the data for this predominint method of listening were compared individually with that of method 1 and of method 4. The results of the analyses are shown in Tablee 47 and 48.

## TABLE 46

ANAISIS OF HOHOABNEITE OF DISTRTBUTION OF SCORSS ACOORDTG TO THE MUTHOD USED PREDOMTNANTLY IN KISTENTMG FOR ACOURACY OF HARMONIC HEFECT

| Scores | Melodic analysis: | Predominant Method <br> Harmonic Style and effect analyais (no technical anal) |  | Totals |
| :---: | :---: | :---: | :---: | :---: |
| $40+$ | 36 (a) | 39 (b) | 5 (c) | 80 |
| 20-39 | 77 (d) | 33 (e) | 15 (i) | 125 |
| -19 | 25 (g) | 7 (h) | 6 (i) | 38 |
| Totals | 138 | 79 | 26 | 24.3 |

$$
x^{2}-15.662
$$

$$
\text { d.I. }=4 \quad P<00
$$

## TABLE 47

 FOR ACCURACX OF HADMONIC EFFEST

| Scores | Predomin <br> 1. Molodio <br> (reading 1 or 2 parts at a time) | Hethod 2. Harmonic (analyzing corplete chord qualetyy | Totals |
| :---: | :---: | :---: | :---: |
| $40+$ | 36 | 39 | 75 |
| 20-39 | 77 | 33 | 110 |
| -19 | 25 | 7 | 32 |
| Tatals | 138 | 79 | 217 |

$$
x^{2}=12.746
$$

d. ${ }^{2}=2$
$P<.01$

TABLIS 48

COMPARISON OF Method 2 am lethod 4 OF LISTENTMA FOR ACCURACY OF HARHOMTC EFFECT

| Scores | Predominant method <br> 2, Marmonic <br> (analyting complete <br> chord quality) | General effect <br> (no technical <br> analysis) | Totals |
| :---: | :---: | :---: | :---: |
| $40+$ | 39 | 5 | 44 |
| $20-39$ | 33 | 25 | 48 |
| -19 | 7 | 6 | 13 |
| Totals | 79 | 26 | 105 |

$x^{2}=8.518$

$$
\text { d.f. }=2 \quad \mathrm{P}<.02
$$

A significant difference was found between the scores of the subjects using method 2 and those using method 1. The difference between the scores of the subjects using method 2 and of those using method 4 was in the region of doubt, but this, for all practical purposes, was also regarded as a dignificant difference, for only a comparatively snall number of subjects had used method 4 predominantily; and with that limited data it could not be confidently concluded that thore was no difference. Although any conclusion drawn from the obtained results had to be tempered by the rather subjective manner of obtaining and classifying the data, it appeared justifiable to conclude from the evidence that the predominant use of method 2 , analysis of complete chord quality, was significantly assooiated with higher scores on the test. It wes interesting to observe, however, that many more subjects used method 1 , reading one or two parts at a time (melodic analysis), than used method 2 , and the smallest number relied only on method 4 , which involved no technical analysis of melodic intervals or harmonies.

Analybis or Recognition of Chord Function

Part $b$ of No. 10 of the quentionnaire provided four categories of response to the question, "to what extent did you analyze the chords or recognize their function as being $I$, IV, V, etc.? ${ }^{4}$ of the four responses only one was to be checked. The responses were: (1) most of the time, (2) frequentiy, (3) only occosionally, (4) never. The responses were tabulated for the four categories, and the distributions of the corresponding test scores were compared. The results are given in Table 49.

Table 49

Anaiysis of homodenity or distribution of scorms ACCORDIMG TO EXTENT OF AMELYSIS OR RECONITION OF CHORD FUNCTION

| Seores | Extent of Host of the time | analysis of Frequently | hord function Occasionally | Never | Totals |
| :---: | :---: | :---: | :---: | :---: | :---: |
| $40 \%$ | 8 | 20 | 43 | 11 | 82 |
| 20-39 | 7 | 29 | 79 | 14 | 129 |
| -19 | 3 | 8 | 22 | 7 | 40 |
| Totals | 18 | 57 | 14.4 | 32 | 251 |

$$
x^{2}-3.376
$$

$$
\text { d.f. }=6 \quad .80>P>.70
$$

The distributions of scoras between the four categories were found to be homogeneous, and therefore there appeared to bo no relation botweon the extent of use of analysis or reoagnition of chord function and achievenont on the tiest. The great majority of the subjects indicated that they analyzed or recognized the function of chords only loccasiona11.t." Those who tudicated that they never analyzed the chords apparently did as well as those who indicated that they analyzed or recognized chord function to some extent.

In the rehoarsal stitnation sumated by the test, with the music moving steadily along; manifestly there le little time in which to anam Iyze or recognize the function of the chords. For this reason, in the move functional courses in ear training end harmony, the recognition of chords as they function in cadences is stresaed primarily, and it was assumed thet whatever uta was made of a knowledge of ohord function, it probably was used most in analyzing the cadences of the phrases. Be that as it mev; the question raised by this study was; how much knowimedge of chord function is actually necessary for successful eohevement on a test which presumably required a functional application of hamonte knowIedge?

## sumbary of Results

The results of the atatistical tests of significance reported in this ohapter are wumarlied as follows:

1. The sampla was found to be homogeneous among the schools for each category of music theory training, and the subjects from all schools in each category wrere pooled.
2. Using the cxiterion group ad each of the groups elassified according to music theory training status, a stratified random sample of fifty test results mas selected for use in calculating the rellability of the test. The reliability of the full length test, estimated from the correlation of the aplit halves based on blternate items, was found to be .98. bufficiently high for good individual prediction. The standard error of measurement was 3.41 .
3. Levele of achtevement were Sound for the masletans in each category of theory tratning by computing the mean scores. For tha group having one year of troining, the mean yas 27.93 for the group having two years, the mean was 38.94 ; for the group having three or more gears, the mean was 44.34.
4. A significont difference was found between the mean score of the group having one year of thoory and that of the group having two years of theory trainingy but no aignificant difference was found between the mean of the latter group and that of the group heving three or more years of theory training. There was, therefore, a significant association betwen the amount of theory training and achievoment on the test, but only up to a point. Havine tro years of training resulted in significantly better achievement, but three or nore years did not result in a further significant increase in the skill measured by the test.
5. The validity of the test was established in three ways:
a. The varlation of the scores between the critexion group, the group having one year of theory tratidngy and the pooled groups having two, and three or more years of theory training, was analyzed by mothod of ranks. The ranic correlation ratio obtained was 61 , indicating that the variation of the score ranks tas substantially in accord with that expected for each level of develop ing and developed musicianship, one evidence of which is manifested in increased competency in the sidll measured by the test.
b. A very vignificant difference wes found betreen the mean score of the criterion group and that of each of the croups classified according to the amount of theory training. The criterion troup, composed of Habove averagel musiciane, a part of their expected genarally superior mulcianship, excelled in that phase of musiw cianship associated with score reeding.
c. A simiflcant associetion was found between school marks In ear-training courses and achievement on the test (see lo. 6 below). This was regarded as the most re1iable indication of valldity insofar as school marks in melc courses can provide satisiactory criterias since ear-training ls most closely aelated to the skill measured by the test-reading, understandings and percelving the melodio and hamonic intervals and chords.
6. The assochation between echool marks in theory courses and test achievement, for the subjects in each category of mount of training, was found to be as follows for the courses included in the studys
e. Eerotraining: significent for each catagory.
b. Harmony in the region of doubt for the group having one year or lass of training; insigmificant for the other groups.
c. Keybond harmonyt significant for the group having one year or loss of tradnings in the region of coubt for the group having two years of training; insignificant for the group having three or more years of training.
d. "Theory", stghifoant for the group having one yoar or less of traintnt (the only group for mich data were available.)
7. The results of the comparisons in tost achtevement, with selected factors or characteristics es variables, were found to be as follows:
w. No relationship was found between age and achzevement on the tost.
b. No relationship was found between sex and achievement on the teet.
c. No rolationship was found between the part usually sung in ensemble and achievement on the test.
d. There was a significant difference bebreen plano majors and majors in other areas of performance in achievement on the test. Hiven between piano majors and voice majors there was a differonce the the . 02 level of confidence, although voice majors would normally use the piano to a considerable extent.
e. The achievement of all subjectis on the best improved significantly only after approxinately six years of piano
expertence, according to the ase data. The amount of piano experience of voice majors had no relation to test achievement up to and including epproximately six years, but there was a difference at the .05 level of confidence for teven of more years of experience. The amount of piano experience of the instrumental majors alone was found to have no association with achievement on the test.
f. No differonce was found in the tast achievement of all subjects according to the mount of formal study on the piano up to and including five yoars of study, but a significant differonce was found for six or more years of study.
g. There was no relation between the number of yearts of formal voice study of volce majors and achlevement on the test.
h. The amount of singing experience of voice majors had no relation to test achievement.
8. There was no difference between string, brassy and woodwind najors in achievement on the test.
J. The anount of study by instrumental majors (strings brass. and woodvind) on their prinolpal Instrument had no rem lationship to test achievement.
k. The emount of experience of instrumental majors on their principal instrument had no relationshitp to test achaveracnt.
9. Possession of the abiluty to read hymon and simple accompaniments at sight on the piano pas found to have a significant association with higher test achievement.
n. The ability to improvise some type of piano accompantment for simple melodes was found to have a aignificant association with higher test achievement at the . 02 level of confidence.
n. With limited datay the amount of ohoral directing experience was found to have no significant assoclation with achievement on the test.
o. With limited data, the amount of instirmental directing experience was found to have no slenificant association with achievement on the test.
p. The Investigation of the method used predominantly in Iistening for accuracy of hamonic effect indicated that the majority of subjects Instened predominantly to one or tro parts at a time, but those who listened for chord quality created by all the parts simultanem ously achieved significantly better on the test. Those who listoned for chord quality did better than those tho listened primarliy for appropriateness of the general effect the 02 level of confldence. These findings mast be tempered by the rather subjective nature of the datag and by the fact that only the precominant, and not the exclusive nethod of Instening used was sought. 具ost subjects indicated that they used two or all three, methode to some extent.
q. Mo relationship was found between the extent of analysis or recogntition of chord function and achievement on the tert. The majority of subjecta analyzed chorcis or recognized their Iunction anly "accasionally." Those who never andyzed on recognized chord function apparently did as well on the test as those who analyzed to some extent.

## Addenda

Thirty-four tesu results from a unversity with a major school of music vere received too late for Inclusion in the study. These returns were tused, however, to provide an estimate of the representativeness of the sample of musiciane and schools obtained for the study. The thirtyfour subjects had all had three or more years of muic theory traning, and the comparison was therefore made with the subjects pooled in that category from the schools nsed in the study* The means and the distribution of scores of the two groups were compared, and the following results were obtained.

The mean score of the sample in the stucty was 44.34 , that of the group of thfirty-four was 43.79. The vardances were found to be homogeneous, and there was no eignificant difference between the means.
(The probability that the difference between the meane was due to chance was between 80 and 90 chences in 100.)

The distribution of scores of the tho groupe was Sound to be very homogeneous. The probability of the observed afficerences occurring by chance wae between 98 and 99 chanoes In 100.

Since the subjects in the late return were so homogeneous in their achievement on the test vith the subjects of the sample, the sample obtained for the study was judged to be quite adequately representative of the musicians and schools for which the test was intended.

## CHAPMER VI

SUREARY, COMCLUSIONS, AMD RECOMAMMATIONS

## Sumary

The purpose of the study mas to measure, in musicians preparing for musical careers, the ability to deteot melodic and harmonic errors in the performance of choral musio while reading the score. Its purpose wes also to detemine levels of achievement according to music theory training status; and to determine the relationship of other selected factors to the ability being measured.

The ability to read a score rapidly and accurately is important for musicians, particulariy for those who find need of it in their daily activitiea. One very inportant use is in the rehearsal situation, where the tones and harmonies, as important olements of music, nust be read and heard correctly in order that the music may be performed as indicated by the score. Avidence of skill in the accurate reading and hearing of tones and harmonies in a practical gituation is therefore provided by the detection of melodic and harmonic errors in the performance of music while inspecting the score.

Despite the recognized necessity of ability in reading the score for well rounded musicianship, the extent of the ability as actually found in masicians has not been well determined. As a review of avail.able music tests revealed, no neasure of the extent of the ability as it actually functions in a musical situation hed been developed. Such a measure appeared to bo urgently needed. It mould help to filll a need
for a measure of achievement in the practical application of the training related to the score reading aspects of musicianship. A knowiedge of some levels of achievement in score reading would be very ugeful in eventually establishing desirable standards of achievement.

Levels of achievement according to the anount of music theory training were sought because trainang in theory, as part of its purpose of developing musioianship, deals largely with roading; understanding; and hearing the structure and function of tones and harmonies. A knowledge of harmony has been found to have one of Its most practioal applications in comparing the sound of hamonies with their notation during the rehearsel of musical groups. Therefore, by seeking such levels of achievement in the sldil, the contribution of theory training to the necessary musical goal of accurate score reading could be determined.

In addition, the relationship of seleoted factors or characteristics, other than thoory training; to score reading ability needed to be investigated. Any knowledge which would contribute to an understanding of the development of such a complex akill as score reading would appear to be welcome.

Among the characteristics which were included in the study were: the princtpal medium of performance, length of study and experience in the principal medium of performance, directing expertence, keyboard facility, age, and sex. Because of the value of the derived information to theory teachers, an attempt was also made to investigate the relationship between achievement in score reading and the method of listening used predominantly in determining the accuracy of harmontc effect, and between score reading achiovement and the extent of analysis or recogntion
of chord function.
To carry out the purpose of the studys it was necessery first to construct a valid and reliable instrument which measured, in a simulated practical musical altuation, the ability to detect discropancies botween the notation and performence of melodic and harmonic intervals and chords. It was then necessary to obtain and interpret data concerning theory training status and other characteristics selected for study. The administration of the instrument to a representative sample of the groupe for which it was intended, 1.a., undergraduate and graduate music najors in oolleges and mivorsities, was a necessary stop in the attain ment of these objectives.

The test was in the choral mediun, because that appeared best suited to the rationale of the study. Hrrors in performance were such as could readily be caused by inaccurate singers during an actual masical rohearsal. But they were introduced in such a mamer as to test the simultaneous correlation of eys and ear in the accurate reading and hearing of notation, and not musical senstivity alone. Since the test was in the chorel medium, and therefore would be of particular interest to choral directors, the musteal passages it contained were relected from a representative semple of readily avalable literature which the average high achool. college, or church choir might use. And becsuse the test was concerned with accurate reading and hearing of tones and harmonies, and could therefore possibly measure the extent of a functional use of knowledge of hemmony, the errors of performance occurred in a representative sampling of the commonly used chords.

In order to test score reading skill as it is found to be useful In a practical gituation, the mutical excerpts selected for use in the test were Ieft mehanged from the composerst Intentlons, and were a phrase or more in length, providing complete contextual meantig for each tone or chord.

The constraction of the test followed recognized procedures. The preliminary tests were andintered to the same type of Individuals and groupe as those for whith the final form was intended. The test items (the errors in perfomance) were endyued for difficulty and discrimfnating power, and those retained in the final form represented the best judgrent of the author in balancing all considerationa involved. the adequacy of the method of marking and scorlng was carefully analyzed. The performances of the nusical excemps, with thoir introduced orrors, were recorded on magnetic tape, using the best singers and equipment avollable. The scores were reproduced legibly, simulating printed music as much as possiole.

To obtain data regarding masie theory training status ond other characteristics, a questionntre was devised to accompany the test. The administration of the questionnatre and test was estinated to require no more than fifty minutes.

The sample obtained for the stady consisted of 260 malcians from oight colleges and univorsities Iocated in five statos. For validation purposes, a criterion group of twenty-five "above average" musicians was obtained. Their eligibility for inclution was deternined by the application of simple oriteria of musical superiority: (1) occupstion of an advanced pasition of musical responsibility, or (2) achievement of
suparior masical resulte.
The statiatical techniques used th the analjses of data were: productmoment coeffloient of correlation, chi-squere, t-tests, f-tests and analysis of veriation method of ranks.

The findings of the study are sumarised as follows:

1. The ramples from each school were homoeneous in the bkill. measured hy the test in each of three catogories of anount of theory traininge (1) one year or loss of trainings, (2) tho years of training and (3) three of mone years of training.
2. The reliability of the test, estinated from the correlation of split halves using alternate items, was Round to be 98 , indicating that tha test wes capable of predicting either Individual or group diffexences.
3. Levels of acheyeming as oxpressed by mean scores, were found for the musticians in each of the above categories, and also for the mugiaians in the amterion group. For the group having one year of tratning, the mean was 27.93 ; for the group having tro years, the mean was 38.94 , for the group having three or more yeatrs, the mean was 44.34 . The mean of the criterion group tas 73.36.
4. A elenifleant difference in teat nchjevenont was found between the group having one year or less of theory training and that having tro years, but no signifleant difference mas found between the lattor group and that having three or more years of training. Having one year or less of training, or havine two jearg, made a significant difference in test achitevement, but heving nore than two years did not appear to contribute significantly further to the skill measured by the test.
5. The test was found to be valid by three methods:
a. A rank correlation ratio of 61 wes found by analysis of the variation of the renired scores of the eriterion growg the group having one year or less of theory training, and the pooled croups having two; and three or more, yeare of training.
b. A tery olgnificent difference mas found botween the mean score of the oriterion group and that of the group in each category of theory troining. The baste astumption was proved that superior musicians, as part of their
generally superior mustcianship, would excel in that phase assoct ated mith score reading.
c. For the subjects in each category of theory tratining, a significant association was found between the most recent school marks in ear-training courses and achievement on the test, Marhs in earmuraining were aosumed to predict more accurately than marks in other theory courses the ablifty of musicians to understand, read, and hear the tones and harmonies of the score.
6. The association between most recent school marks in theory courses and test achievement, for the subjects in each catagory of theory training, ras found to be as follows for the courees included in the studye.
a. Ear-tratning" signtificont for each category.
b. Harmony: of doubtul significance for the group having one year or less of trainingy insignisicant for the other Eroups.
e. Reyboard harmonyt significant for the group having one year or less of trainingy ponsibly significant for the group having two years of tratningt insigntficant for the group having three or more years of training.
d. "Theory": significant for the group heving one year or less of trating (the only group for which data were avallable).
7. The following factors were found to have a significant relam tionship with hisher test achievoment:
a. The choioe of pano or organ as the principel medim of performance.
b. Having approximately stx years or nore of piano experLence (significent for ell subjects pooledy possibly significant for voice majors baving approximately beven or more yeark of experience; mot signtificant for string, yoodrind, or brass majors alone).
c. Heving approximately six yoars or more of formal study on the pieno.
d. The ability to play hyms and stmole accomanimonta at sight on the piano.
e. Possibly the ability to improvise sane type of piano accompaniment for simple melodies.
f. The predominent use of tistentice for the mode or quatty of chords (najory ninor, etc.) in determining the accuracy of harmonic effect. The majortty of subjects, however, used predominarity a nelodic approath, reading nad lisueang to one or two ports at a time, in. determining the accuracy of performence.
8. The following factore selected for study were found to have no significant reletiton to achieverent in the skill reasured by the test:
a. Age.
b. Sex.
c. Parb sumg in ensemble.
d. The mount of singing expemience of formal roice ebudy of volce mejors.
e. The choice of a string, bress, of woodrind Instrument as the principal mediun of performance of inctrmental majors.
f. The amount of experience or formal study ky instrumental majors (atring, breas, or woodrind) on their principal Instromate:
g. The amount of oither ohoral or ingtrumental directing expertence (from the data rvailable).
9. The extent of anelyeis or recoguition of chord function (as chorás balng I, IV, T, etc.).

## Bonclusions

From the results of the study the following conclustons vere reached:

1. A valid ard reliable test wat constructed to measure the ability of rusicians to detect nolodic and hamonic exrors in the performance of ctioral nusic while inspacting the score.
2. The test is unique in measuring that aspect of score reading ability concerned with the aural-visual discrimination of tones and hamoniee as they occur in the context of complete musleal phrases, and as that ability functions in the practicel situation of a choral rehearsal.
3. The test points the way tovart fulfilling a need on the undergraduate and traduate levels for a measure of the practical application of the training related to the score reading aspects of musicianship, or for a measure of the functional value of the training which is designed to contribute to score reading ability.
4. According to differences in levels of achievement, having one year or less, or two years of theory tradning made a significant difference in achiovement in the blill measured by the test. Thus, the courses in theory offered in the first tuo years of the tusic training program, in their cumplative effect, are functional in the mplication of the training to a practical musical situation. Having more than two years of theory training did not result in a significant increase in the skill measured by the test; therefore, the edvanced, or upper level courses do not appear to be of practical use with regard to this skill.
5. A keyboard instrument is the best background for higher achievement in score reading. Hovevery it is necessary to have experience or formal study on the keyboard for a period of approximately sixx years before score reading as neasured by this test is significantiy improved. and ability in sight reading material at the piano should be developed at least to the point where hyms and exsy accompaniments can be played is score reading is to be olignificantly improved.
6. Reading and listening to all parts simultaneously so as to recognize the quality (mode) of ehords and harmonic interwals ahould be the method predominantly used in determining the accuracy of their performance. This does not imply, however, that reading and hearing one or two parts at a time should not be used to some extent.

After the statemont of these conclusions it seems appropriate to offer the following observation, dram from an anolysis of the data and results of the study. Some of the results regarding the factors which were found to have a significant association with test achiovemont were contrary to what could be expected. One would ordinarily expect increased experience with the materials of music to be associated with increased achievenent in score readings but age and the anount of dirocting experience were found to have no significant relationshtp with
the ability measured by the test. Also, upper level and graduate stum dents, who had more formal training in music courses and more expemence with male, did not achieve significantly better scores than students with only tro years of acadenic training. Some few individuals with comparatively littlo acaderic music traintig scored as high on the test as the highest scoring menbers of the oriterion group. The inference may be made that, wis with so many other attributes of musicianship, abilIty in the arral-visual aiscrimination of tones and hamontes is associated with the natural endoments of the individual and that the dbility is developed earlier in those individuals having the greatest anount of such endoments.

## Recommendations

The resulte of the study lead to the following recomendations for music education programs on the college level:

1. The finding of signiftcant difference in the ability measured by the test between musicians having at loast tro years of theory treining and those having only one year or lass sugsests that two years of theory trainine should be the minimm for music students who need to acquite some competence in score reading, or in activities Involving the reading of scores.
2. For Increased compatence in reading accurately, comparatively simple scores such as were used in the test, it is necessary to have increased experience on a leyboard instrunent. Since only after a period of years-from the date of this thady, approximately six years-does the cumitative effect manifest itself in increased score reading competence, it 13 advisable that all music majors should have as much piano experience as possible if it is desired that score reading ability bo signicicantily improved.
3. Ability at the keyboard should be developed at least to the point where material such as rum tunes and easy accompaniments can be played at sight. And possibly the ability to

Improvise at the piano should be developed to the point where sccompaniments can be provided for aimple melodies.
4. If it is destred thet achievenent in theory coursos be meacured in teme of functional shills acquired, among which would be the abillty to road and hear the tonal content of nusic, the testing technique utilized in this study could be employed.

Other recomendations, and suggestions for further research; aristig from the study ere offered as follows:

1. The provision of nome for this test, or a similar one mould malce it more nisefill for purpoees of comparisons betreen individuals, groups, and schools.
2. The choral mediun was selected for use in this study but by Its nature it was not the clearest medium in which to detect deviations in performence from the notation. The entire strudy could profitably be repeated, usting some other medium, such as plano, organ, or some combination of instruments, provided that they met the requirements of the rationale of the study.
3. The procedures and tecimiques employed in this study could be applied to measuring functional skili in reading and hearing the riythmic elements of musle.
4. A stuay could well be made to deternine what lenoriedge of the functional interrelathonships of chords is actually necessary for musiciana in order to read and hear the score accurately. Such definite knomledge would be of value in detemfning the content and method of courses in harmony, ear-trainingy and harmonic analysis.
5. The ability to read and hear eccurately the melodic intervals of single tine melodies in a functional situation is obviously important. Fior instance, the teacher of voice or instrunents needs to be able to detect and correct all melodic errors. A measure of this ability would be very usefuly, but as yet none ls dxaileble.
6. An interesting and worthrile study would be to analye the effect of context upon the perception of either or both the quality (major, minor, etc.) or the function (tonic, dominant, etc.) of chords.
7. The effect of various tone qualities upon the perception of chord quality could also be investigated. For instance, the same musical passage, containing introduced errors of
> performance, could be periomed by volces, etrings, woodminds, brasses, or combinations of these, and their influence upon the acouracy of eyemetr perception of the tonal elements could be investigated.

## Concluding Stavement

It is apparent that the ablitity to read a nualcal score accurately is dopendent not only upon the innate masical potentialities of the individual, but anso upon a spectife lind of training.

Coumser in music theory should be specificolly designed to atd the student in devoloping a keen aural awareness of the meaning of maical symbols. It is aigatricant that there is listle improvement in the ability to detect errors in the performance of a score after the second Jear of college training. Altwough there is the possibility that the loreling off of this sbility is due to other causes, it is equaliy probable that musied training ator the secone year in college does Little to devolop the abllety in question.
of equal Importance is the rolating of the pymbolism of a score to a space-frame. The keyboard (piano or organ), unlike monophonic instrum ments, serves as a means of reproduche several parts or broteest shmultaneously and thm becomes extremely inportant in aiding the learner to develop the ability to read and hear musical concords accurately.

It would seem thet a well developed facility at the keyboard not only permts the learner to reproduce accurately the sounds symolized In a masical score but, over a poriod of time (apparentiy at least six Jears), develops an aural and visual amareness of particular combinations or configurations of tone.

It appears to be sienificant that those who were the nobt succassful subjects in this stady stated that thein primary check on the accuracy with which the woch score was sung was first to Itsten to the quality of the chorde. Chordal qualities thus are perceived as entities, and long association with a keyboard undoubtedly helps the musical student to recognize, with a highec Iovel of accuracy, the degree of correspondence betreen what is mritten and what is pexformed.

BIBIIOMRAPMY

Adkins, Dorothy and others. Construction and Analysis of Achievement Tests. United States Government Prining Office, washingtion D. C. 1947.

Agnens Marie. "A Comparison of Auditory Imagee of Lusioiang, Psyohologists, and Ohildrene" Psychological Monographs, 31, Ho. $1_{1}$ whole No. $140,279-287.1922$.

- "The Auditory Inagery of Great Composers ${ }^{\text {" }}$ " Pbychological Monow graphs, 31, Ho. 1, whole NO. 140, 279-287. 1922.

Aliferis, James, Aliferks Haic Achievement Teat. University of winnesota Pressy kinneapolis. 1954.

Aliferis, James and J. E. SteckleIn. Whe Development of a College Entrance Test in Tusic Achievement:" Journal of Hesearch in Music Education, 1, No, 2, 83-96. Fall 1953.

Allen, Richnrd D., Walter H. Butterifeld, and Marguerite Tully. Providence Inventory Test in Musicy Marnal of pirections. World Book Co.s Yonkers, N. Y. 1932.

Anderson, Kenneth E.y E. Thayer Gaston, and Melvin Zack. A Method of Establishing Test Validity Unpublished paper; The University of Kansas. 1954.

Beach, Frank A., and H. E. Schramel. Beach Music Testa; Manual of Directions. Bureau of Educational teasurements; Kaneas State Teachers College, Emporia, Kansas. 1920, revised 1930.

Budge, Helen. A Study of Chord Frequencies Based on the kusio of Representative Composers of the Eighteenth and Nineteenth Centuries. Bureau of publications, Teachers College, Colmbia Oniverbity; New Tork. 1943.

Burns, Samel T. Harmonic Skills Used by Belected High School Choral Leaders. Bureau of Publications, Teachers College, Columbia University, New Iork. 1945.

Davis, Frederick B. "Item Selection Techniques." Educational Measurement. E. F. Lindquist, editor. American Council on Education, Washington, D. C. 1951. 267-285.

Diller, Angela. First Theory Book. G. Schirmer, Incet New Tork. 1921.
Drake, R. M. "Four New Tests of Musical Talent." Journal of Applied Psychology, 17, 136-147. 1933.

[^33]$\qquad$ - Drake Musical Momory Test; a Test of Kusical Talent. Public School Publishing Company, Bloomington, Ill. 1934.

Dykema, Peter W. and Karl Gehrkens. The Teaching and Administration of High School Kusic. O. C. Birchard and Company, Boston. 1951.

Famsworth, P. R. "An Historical, Oritioal, and Experimental Study of the Seashore-Kralwasser Test Battery," Genetic Psychology Monographs, 9, 291-393. 1931.

- "Studies in the Psychology of Tone." Genette Psychology Konographs, 15, 1, 1-91. 1934.

Freeman, Frank S. Theory and Practice of Psychological Testing. Henry Holt and Company, Inc. New Iork. 1950.

Gaston, E. Thayer. A Test of Musicallty Manual of Directions. Third edition, revised. Streep Kusic Company, Inc., Kansas City, Ko. 1950.

Gaw, Bsther Allen. uTive Studies in the Husic Tests." Psychological Monographs, 39, 2, Whole Mo. 178, 145-156. 1928.

Gehricenss Karl Mo Heundamental Princlples of Theory Teaching." yrua Proceedings, 34, 64-71. 1939.

Gildersleeve, Glenn and thaye Soper. Kugical Achievement Test. Bureau of Fublications, Teachers College, Columbia lniversity, New York. 1933.

Greene, Edward B. Measurements of Fuman Behavior. The Odyssey Press, New Yoric. 1952.

Guilford, J. P. Fundamental Statistics in Psychology and Education. MeGrar-Hill Book Company, New Yoric. 1950.

Henarickson, Gordon and Olara G. Stratemeyer. "Eusic Education." Encyclopaedia of ECucational Hesearch, Walter S. Monroe, edftor. The Hacmilen Company, New York. 1950. 761-771.

Hevner, Kate and John Lansbury. Oregon Masical Discrimination Test. C. 1F. Stoelting Company; chicago. 1935.

Hevner, Kate. "Tests for Aesthetic Appreciation in the Field of yusic." Joumal of Applied PsycholoEy, 14, 470-477. 1930.
$\qquad$ - "A Study of Tests for the Appreciation of 筑usic." Journal of Applied Psychology; 15,575-583. 1931.

Hillbrand, E. K. Hillbrand Sleht-Singing Test. World Book Company, Yonkers, Mew York. 1923.

Hogsard，Lara Q．Improving Masic Reading in the Choral Rehearsal．The Author，Hew York．1947．

Hutchinson，Herbert．Hutchinson Husie Tests．Pubile Sohool Publishing Company，Bloomington，I11． 1924.

Johnson，Palner O．Statistical Methods in Research．Frenticemflall， Inces．Hew York＊39149．

Jones，Vincent．Musio Education in the College．O．C．Birchard and Company，Boston． 1949.

Jurgensen，C．D．＂Table for Determining Thi Coefficients．＂Psychometrika， 12；1，17－29．Harch 1947．

Hotick，H．Lela and T．It Torgerson．Diagnostic Tosts of Achievement in $\frac{\text { Musics }}{1950 .}$ hamal of Directions．Calffornia Test Bureau，Los Angeles．

Knuth，willian t．Knuth Aohievement Tests in husic．Educational Test Bureaty Minneapolis．1936．

Kwalwasser，Jacob and G．N．Huch．Kwalwasser－Ruch Test of Musical Accomplishment：Bureau of Educational Research，State University of Iowas Iowa City． 1924.

Kralwassery J．Kwalwasser Test of welodie and Hamonio Sensitivity． Victor Telking Machine Company，Camden，N．J．1926．
＊Kwalwasser Test of Kusical Information and Appreciation． Bureau of Educational Research，State University of lowa；Iowa City． 1927.
＊＂Tests and Measurements in 解sic．＂Psychological Bulletin， 25．284－301． 1928.

Kwalwasser，Jacob and Peter W．Dykema．K－D Music Tests；Manual of Directions．Oarl Fischer，Inces New York． 1930.

Lindquist，E．F．＂Preliminary Considerations in Objective Test Con－ struction．＂Educational Measurement．E．F．Lindquist，editor． American Council on Education，washington，D．G．1951．119－158．

Lowery，H．NCadence and Phrase Tests in Iusic．＂British Journgl of Prychology，17，111－118．1926．

Lundin，Robert 脚．＂The Development and Validation of a Set of lusical Ability Tests．＂Psychological Monographs，63：10，Thole No． 305. 1－20．1949．
－An Objective Psyohology of 解sic．The Ronald Press Company， New York． 1953.

Wadison, T. H. "Tnterval Discrimination as a Haasure of Mustoal Aptitude." Archives of Psychology, No. 206, 1-99. 1942.

WoHose, Allen Irvine. The Contrapuntal Harmonic Technique of the 18 th Gentury. F. S. Crofts and Company, New York. 1947.

Melchery Robert A. THe Status and Function of Ear Training in American Colleges and Gonservatories." MNA Proceedinge, 33, 185-194. 1938.

Korgany Nazel Nohavec, editor. Nusic foucation Source Book. Music Educators National Conference, Chicaso. 1947.

Wosher, Haymon M. A Stucy of the Group Mothod of Measurement of SightAhncing. Aureau of Publicationsy Teachers College, Columbia University, New Tork. 1925.

Aurply: Foward Ansley and Edrin John Stringhan. Creative Harmony and䓝asicianship. Prentice-IIall, Inc.g New York. 1951.
tursell. James I. The Psychology of Music. W. Norton and Company, New York. 1937.

Fursell, James I. and Wabelle Glem. The Psychology of School Music Teaching. Silver Burdett Compay, New Zork. 1938.

Otterstein, Adolph and Raymond Hosher OMA Siph-Singing Test. Stanford University Prese, Palo Alto, Calif. 1932.

Pankrati, Waldo 0. MA Comparison of Some Group Musical Achievement Tests as Related to I. Q. and Teaoher Rating." Umpublishod Master's thesis, the University of Kansas. 1950.

Potran, I. A. "An Analysis of the Ortman Tests in Tonal Discrimination and Memory;" Bulletin of the Anemican Masicological Society, 25, 2, 304-306. 1937.

Pliston, Fralter. Hampry W. W. Worton and Compary, Inc., New York. 1948.

Ross; C. C. Maasurement in Today's Gehools. Prenticemall, Incos Ner York. 2948.

Saetreit, J. Co, D. Levisy and C. F. Seashore. Revision of the Seashore Yeasurea of Musical Talent. State Universiby of Iowa, Towa City. 1940.

Sallsbury, F. S. and H. A. Smith. MPrognosis of Sight-Singing Ability of Normal Sohool Students," Journal of Applied Psychology, 13, 425-439. 1929.

Schoen; Max. "Tests of Mustcal Feeling and Husical Undarstanding." Joumal of Comparative Psychology, 5, 31-52. 1925.

- The Psycholosy of husic. The Ronald Press Company, Now York.

Seashore, C. E., D. Lexis, and J. G. Seetreit. Seashore Moasures of Musicel Talents, revised edition. Educational Department, RCA Victor Division, Redio Corporation of America, Camden, N.J. 1939.

Seashore, C. H. WThe Peychology of Music: XI." Kusic Educators Journa1, $24,3,25-26.1937$.

Sessions, Roger. The Misical Experience of Composery Performer, Histener. Princeton University Presas Princeton, M. J. 1950.

Torgerson, T. L. and Ernest Tahnestock. Torgerson-Fahnestock Masie Lests. Public School Publishing Company, Bloomington, IIl. 1926.

Traxier, Arthur E. MAdrainistering and Scoxing the Objective Test." Educational Measurement, E. F. Tindquist, editor. American Council on Education; Fashington, D. C. $1951.329-116$.

Halker, HeIen M. and Joseph Lev. Statistical Inference. Henry HoIt and Company, New York. 1953.

Whistler, Harvey S. and Louts P. Thorpe. Mustical Aptitude Test Series A. Galifornia Test Bureau, Los Angeles. 1950.

Wright, Frances. "The Correlation Between Achievement and Capacity in Ifusic." Joumal of Educational Research, 17, 50-56. 1928.

APPENDTA A

Original Data


Explanation of colum headings and symbols used in this and the next seven tablest Theory Trafning Category: 1 , one year or less; 2 , two years; 3 , three or more years. Voice Part: $S$, soprano; $A_{2}$ alto; T, tenor; $B$, bass. Principal Performing Mediunt $P$, piano, 0 , organ; $V$, voice; $S$, stringed instrum ment; W, woodrind instmment, B, brass instrument. Play hyms (at sight on the piano)?, and Improvise (plano accompaniments for simple melodien)?: Y, yes; $N_{\text {, }}$ no. Achievement (reported, according to most recent mark), and Actual Achievement (according to school records): 1, above average; 2, average; 3, belon average. (The data for Actual Achievement are provided only for the subjects in this table and in Table 57.) Predominant kothod of listening (for accuracy of performance) 1 , listening to one or two parts at a time; 2, to complete chord quality; 4 , to general style and effect. Analysis or Recognition of Chord Functionz 1 , most of the time; 2, frequently; 3, occasionally; 4, never.

TABLE 50 （continued）

|  |  |  |  |  |  |  |  | Mad＊xedxa＊sx | Trs. Piano Study |  | $\begin{gathered} \text { 買 } \\ \text { 总 } \end{gathered}$ |  |  |  | $\begin{aligned} & \text { Achieve } \\ & \text { mont } \\ & \hline \end{aligned}$ |  |  | Actual Achieve． |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  |  | 鄀 | O |  |  |  |  |  |  |  |  |  |  |  | $\begin{aligned} & \text { 总 } \\ & \text { 荷 } \\ & \text { 采 } \end{aligned}$ |  |  |  |  |  |  |  |
| 6 | 19 | M | 21 | 2 | 4 | V | 9 | 9 | 3 | 3 | － | I | 2 | $\square$ | 1 | 3 | 2 | 1 | 3 | 2 | 2 | 2 |
| 7 | 34 | 喍 | 21. | 2 | T | B | 4 | 4 | 5 | 5 | Y | 7 | 1／2 | 1／2 | 2 | 3. | 1 | 2 | 3 | 1 | 4 | 3 |
| 8 | 8 | F | 21 | 3 | A | V | 4 | 8 | 3 | 8 | Y | － | 0 | $-$ | 2 | 3 | 2 | 2 | 3 | 2 | 1 | 3 |
| 9 | 47 | H | 24 | 3 | B | P | 14 | 17 | 14 | 17 | T | I | $-$ | $\square$ | 1 | 1 | 1 | 1 | 1. | 1 | － | 3 |
| 10 | 62 | 1 | 20 | 2 | B | $B$ | 9 | 9 | 1 | 1 | H | \＃ | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 2 |
| 11 | 44 | F | 18 | 1 | A | W | 9 | 9 | 10 | 10 | 7 | Y | 4 | － | 1 | 1 | 1 | 1 | 1 | 1. | 1. | 3 |
| 12 | 30 | \％ | 20 | 2 | T | \％ | 1 | 4 | 2 | 2 | N | F | 1. | 4 | 2 | 3 | 2 | 2 | 3 | 2 | － | 2 |
| 13 | 51 | \％ | 24 | 3 | B | 핲 | 3 | 10 | 3 | 3 | I | H | － | － | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 2 |
| 14 | 46 | F | 22 | 2 | A | F | 15 | 15 | 15 | 15 | Y | Y | 0 | － | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 2 |
| 15 | 29 | F | 19 | 2 | 5 | $\nabla$ | 2 | 6 | 7 | 7 | Y | Y | 0 | － | 2 | 1 | 1 | 1 | 1. | 1 | 2 | 3 |
| 16 | 45 | F | 20 | 3 | s | $p$ | 11 | 11 | 11 | 17 | 7 | $\Psi$ | 0 | － | 1. | 1 | 1 | 1 | 1. | 1 | 2 | 3. |
| 17 | 61 | F | 20 | 2 | 5 | V | 2 | 4 | 4 | 4 | I | I | － | － | 3 | 1 | 2 | 3 | 1 | 2 | 1 | 3 |
| 18 | 31 | F | 19 | 2 | A | Wi | 8 | 8 | 13 | 14 | I | $\pm$ | 1 | 3 | 7 | － | － | $\cdots$ |  |  | 2 | 3 |
| 19 | 27 | 1 | 20 | 2 | B | P | 10 | 12 | 10 | 12 | $\underline{Y}$ | I | 1 | 1 | 2 | 2 | 2 | 2 | 3 | 2 | 4 | 3 |
| 20 | 20. | F | 21. | 3 | \％ | V | 4 | 7 | 6 | 8 | $\Psi$ | I | 2 | － | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 3 |

H


|  |  |  |  |  | $\begin{aligned} & \text { 曹 } \\ & \text { H } \\ & 0 \\ & \hline 0 \end{aligned}$ |  | $\text { Hde Apngs } \cdot \operatorname{six}$ |  |  | $\frac{6}{\text { 爰 }}$$\frac{8}{8}$$\frac{8}{8}$ |  |  |  |  | Achieve nent |  |  | Actual Achieve． |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| $\begin{aligned} & \text { 8 } \\ & \frac{0}{7} \\ & \frac{م}{\circ} \end{aligned}$ |  | 菖 | $\Phi_{0}^{9}$ |  |  |  |  |  |  |  |  |  |  |  |  | $\begin{aligned} & \text { 㱐 } \\ & \text { 总 } \\ & 4 \\ & 4 \end{aligned}$ |  | $\begin{aligned} & \text { 右 } \\ & \hline \end{aligned}$ | $\begin{aligned} & \frac{0}{4} \\ & \frac{4}{4} \\ & \frac{4}{4} \\ & 8 \\ & 8 \end{aligned}$ |  |  |  |
| 36 | 19 | T | 18 | 1 | A | $p$ | 5 | 9 | 5 | 9 | I | $\pm$ | 0 | － | 2 | 2 | 2 | 2 | 2 | 2 | 1 | 4 |
| 37 | 42 | H | 20 | 3 | T | V | 3 | 11 | 6 | 10 | H | I | 0 | － | 2 | 1 | 1 | 2 | 2 | 1 | 1 | 3 |
| 38 | 19 | 3 | 20 | 2 | H | \＄ | 5 | 10 | 2 | 2 | N | H | 2 | 2 | 2 | 2 | 2 | 2 | 1 | 1 | 2 | 1 |
| 39 | 75 | 典 | 21 | 2 | B | P | 10 | 10 | 10 | 10 | T | Y | 7 | 1 | 1 | 1 | 1 | 2 | 1 | 1 | 4 | 4 |
| 40 | 39 | F | 19 | 2 | A | 9 | 11 | 12 | 11. | 12 | T | 7 | － | － | 2 | 2 | 2 | 1 | 1 | 1 | 1 |  |
| 11 | 17. | H | 18 | 1 | B | B | 2 | 6 | 2 | 2 | H | N | 0 | 0 | 1 | 3 | 2 | 1 | 2 | 1 | 1 | 4 |
| 42 | 100 | F | 21 | 2 | A | P | 9 | － | 9 | $-$ | T | I | $\underline{-}$ | － | 1 | 1 | 1 | 1 | 1 | 1 | 1 |  |
| 43 | 24 | F | 23 | 3 | A | 5 | 4 | 0 | 9 | 12 | I | I | 重 | 0 | 2 | 2 | 2 | 2 | 2 | 2 | 1 | － |
| 44. | 33 | F | 18 | 1 | A | $p$ | 11 | 11 | 11. | 11 | T | 7 | － | － | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 3 |
| 45 | 54 | F | 39 | 1 | A | P | 13 | － | 13 |  | I | － | － | － | 2 | 2 | 1 | 1 | 2 | 1. | 2 | 2 |
| 1.6 | 59 | 花 | 24 | 3 | B | P | 10 | 10 | 10 | 10 | $\Psi$ | $\pm$ | － | 2 | 2 | 2 | 1 | 1 | 1 | 1 | － | 4 |
| 47 | 24 | F | 18 | 1 | A． | $\nabla$ | 1 | 5 | 8 | 8 | Y | H | 0 | － | 2 | 2 | 2 | 1 | 2 | 2 | 2 | 3 |
| 48 | 4 | F | 21 | 3 | S | $p$ | 15 | 15 | 15 | 15 | Y | I | 1 | － | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 2 |
| 49 | 56 | T | 20 | 3 | A | $p$ | 14 | 15 | 14 | 15 | $\Psi$ | Y | － | － | 2 | 1 | 2 | 1 | 1 | 1 | 2 | 4 |
| 50 | 40 | F | 20 | 2 | 8 | F | 12 | 12 | 12 | 12 | $\pm$ | T | 1 | 1 | 2 | 1 | 2 | 2 | 1 | 2 | 1 | 4 |

```
TADLE 50 (continued)
```



```
TABLE 50 （continued）
```

| $\begin{aligned} & \text { 世 } \\ & \stackrel{0}{0} \\ & \stackrel{\rightharpoonup}{2} \end{aligned}$ |  | $\begin{aligned} & \text {-6 } \\ & \hline \end{aligned}$ | 8084 |  |  |  |  |  | Frs．Piano Study |  |  |  |  |  | Ac <br> 膏 眮 | ieve ent <br>  |  | Ac Ac B 㗊 | tual <br> ieve <br>  | 古 离 容 |  | $\begin{aligned} & \text { 4. } \\ & 0.0 \\ & 0 \\ & 0 \\ & 0 \\ & 0 \\ & 0 \\ & 0 \end{aligned}$ |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 66 | 33 | F | 19 | 2 | S | P | 11 | 11 | 11 | 11 | I | $\mathbf{T}$ | 0 | － | 1. | 1 | 1 | 1 | 1 | 1 | 2 | 2 |
| 67 | 37 | y | － | 2 | \＄ | B | 6 | 7 | 4 | 4 | T | I | 0 | 0 | 1 | 3 | 2 | 1 | 3 | 2 | 2 | 4 |
| 68 | 66 | F | － | 1 | 溒 | P | 13 | 13 | 13 | 13 | Y | I | 0 | $\sim$ | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 2 |
| 69 | 31 | H | 25 | 1 | B | V | 1 | 1 | 4 | 8 | 7 | H | 0 | － | 2 | 3 | 2 | － | $\pm$ | － | 4 | 4 |
| 70 | 27 | F | 21 | 3 | A | $P$ | 12 | － | 12 | － | 7 | $\pm$ | － | － | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 3 |
| 71 | 40 | M | 18 | 1 | \％ | 0 | 2 | 5 |  | 12 | $\pm$ | I | 6 | － | 2 | 2 | 1 | 2 | 1 | 1 | 1 | 2 |
| 72 | 22 | F | 19 | 1 | S | P | 5 | 5 | 52 | 5 | I | T | 0 | － | 2 | 2 | 2 | 2 | 2 | 1 | 1 | 3 |
| 73 | 26 | 者 | 19 | 2 | B | B | 6 | 6 | 4 | 5 | 7 | I | － | 0 | 2 | 2 | 2 | 2 | 2 | 2 | － | 3 |
| 74 | 45 | F | 20 | 2 | S | F | 8 | 8 | 8 | 8 | I | $\pm$ | 3 | － | 2 | 1 | 1 | 2 | 1 | 1 | 2 | 2 |
| 75 | 21. | H | 19 | 1 | B | \％ | 4 | 4 | 1 | 2 | H | I | 0 | 0 | 1 | 1 | 2 | 1. | 1 | 2 | 1 | 3 |
| 76 | 83 | F | 20 | 2 | A | $P$ | 8 | 74 | 8 | 14 | $\pm$ | Y | 考 | － | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 2 |
| 77 | 1.4 | \％ | 20 | 2 | T | P | 10 | 10 | 10 | 10 | Y | H | 3 | 2 | 1 | 2 | 1 | 1 | 2 | 1 | 1 | 1 |
| 78 | 31 | H | － | 3 | $T$ | B | 9 | 9 | 2 | － | $\mathbf{T}$ | H | 3 | 10 | 2 | 1 | 3 | 2 | 2 | 2 | 1 | 3 |
| 79 | 52 | F | 20 | 2 | A | V | 27 | 8 | 14 | 14 | I | 7 | 1 | $\underline{-}$ | 2 | 1 | 2 | 1 | 1 | 1 | 1 | 3 |
| 80 | 16 | F | 21 | 3 | A | P | 10. | 10 | 10 | 10 | F | I | － | － | 2 | 3 | 2 | 1 | 3 | 1 | 4 | 3 |



## MABLS 51

DATA FOR MUSICIANS DI SCHOOL B


```
TABIE 51 （continued）
```

|  | 9 0 0 0 4 0 0 | \％ | 8 |  |  |  | Yrse. StuAy PRM | Mide rouxa－日x | SOMAS out fa esx | Tro．Pizno Exper． |  |  |  |  | Ach $\square$ |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 11 | 27 | H | 18 | 1 | B | 5 | 6 | 6 | 1. | 2 | $\pm$ | H | － | － | $\cdots$ | $-2$ | 1 | 2 |
| 12 | 21. | H | 19. | 2 | B | P | 8 | 8 | 6 | 8 | Y | Y | 0 | － | 2 | 32 | － | 3 |
| 13 | 22 | F | 18 | 1 | A | $\underline{F}$ | 10 | 10 | 10 | 10 | I | － | 0 | － | $-$ | $-2$ | 2 | 4 |
| 14 | 26 | F | 19. | 2 | A | V | 5. | 6. | 3 | 4 | $\underline{T}$ | Y | 0 | － | 2 | 32 | 1 | 4 |
| 15 | 40 | F | 19. | 2 | S | $\nabla$ |  | 10 | 2 | 2 | I | $\pm$ | － | － | 2 | 22 | 2 | 3 |
| 16 | 36 | $F$ | 19 | 1 | A | \％ | 6 | 6 | 7 | 12 | $\Psi$ | $\mathbf{Y}$ | 5 | 6 | － | $-1$ | 1 | 2 |
| 17 | 23. | H | 19 | 1 | 7 | B | 4 | 4 | 2 | 2 | $\psi$ | $\pm$ |  | － | $\bar{\square}$ | －2 | 2 | 2 |
| 18 | 53 | 避 | 20 | 2 | T | 8 | 6 | 17 | 4 | 4 | I | V | 4 | 2 | 2 | 22 | 1 | 3 |
| 19 | 20 | H | 19 | 1 | 1 | $\nabla$ | 3 | 5 | 1 | 1 | I | \％ | 1 | － |  |  | ＋ | 3 |
| 20 | 16 | $F$ | 18 | 1 | S | $\nabla$ | 3 | 4 | $1{ }^{1}$ |  | I | W | － | － | － | 2 | 1 | 3 |
| 21 | 28 | 㖵 | 19 | 1 | T | V | 10 | 10 |  | － | \＃ | N | － | － | － | $-3$ | 1 | 3 |
| 22 | 18 | $F$ | 20 | 1 | S | 1 | 4 | － | 4 | － | $\underline{T}$ | － | T | T | － | $-3$ | 2 | 2 |
| 23 | 29 | M | 21 | 3 | B | B | 10 | 12 | 6 | 12 | I | 7 | 1 | 1 | 1 | 1 | 2 | 3 |
| 24 | 24 | F | 18 | 1 | A | P | 10 | 10 | 10 | 10 | $\underline{T}$ | N | 0 | － | － | $-1$ | 1 | 3 |
| 25 | 51 | F | 19 | 3 | A | P | 8 | － | 8 | － | I | $\Psi$ | － |  |  | 1 | 1 | 3 |


|  |  | 荌 | $\stackrel{1}{4}$ |  |  |  | $\begin{aligned} & \text { 氙 } \\ & \text { 总 } \\ & \text { b } \\ & \text { 中 } \end{aligned}$ |  |  |  |  | $\begin{aligned} & \text { 世 } \\ & \text { 害 } \\ & \text { H } \\ & \text { H } \end{aligned}$ |  |  |  | hieve ment <br>  | $\begin{aligned} & \text { 荷 } \\ & \text { B } \end{aligned}$ |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 26 | 29 | H | 18 | 2 | B | P | 3 | 5 | 3 | 5 | T | I | － | $\underline{-}$ | 2 | 2 | － | 2 | 2 |
| 27 | 52 | \％ | 20 | 2 | B | P | 13 | 6 | $1{ }^{1}$ | 6 | N | T | 0 | $\underline{L}$ | 2 | 2 | 2 | 1 | 3 |
| 28 | 94 | F | 20 | 3 | 4 | V | 2 | 6 | 10 | － | \＄ | $\pm$ | 0 | － | － | － | 1 | 1 | 1 |
| 29 | 36 | F | 18 | 1 | 4 | P | 12 | 12 | 12 | 12 | $\pm$ | 4 | 0 | － | － | － | 1 | 1 | 3 |
| 30 | 22 | F | 24 | 3 | A | P | 3 | － | 3 | － | $\pm$ | $\pm$ | － | － | 1 | 2 | 1 | 1 | 3 |
| 31 | 21 | W | 31. | 3 | B | V | 2 | 3 | 2 | － | 7 | \＄ | 0 | － | － | － | － | 1 | 3 |
| 32 | 41 | 数 | 19 | 2 | B | B | 2 | 6 | 1 | 1 | 7 | I | － | － | 2 | 2 | 2 | 2 | 2 |
| 33 | 45 | F | 19 | 3 | A | P | 13 | － | 13 | $-$ | $\pm$ | 7 | － | － | － | $\cdots$ | － | 1 | 3 |
| 34 | 54 | 7 | 19 | 2 | A | P | 12 | 13 | 12 | 13 | I | I | 0 | － | 1 | 1 | 1 | 2 | 3 |
| 35 | 57 | F | 20 | 3 | S | P | 13 | 13 | 13 | 13 | $\Psi$ | T | 0 | － | $-$ | － | 1 | 2 | 1 |
| 36 | 47 | F | 20 | 3 | $s$ | P | 14 | 34 | If | 34 | I | \％ | 0 | 0 | － | － | 1 | 2 | 1 |
| 37 | 32 | F | 18 | 1 | － | P | － | － | － | － | I | I | － | － | － | － | － | 2 | 2 |
| 38 | 13 | $F$ | 19 | 1 | S | P | 3 | 4 | 3 | 4 | \％ | － | － | － | － | － | 2 | 4 | 4 |
| 39 | 25 | M | 18 | 1 | B | B | 10 | 10 | 13 | 23 | 7 | H | － | － | － | $-$ | 2 | $-$ | － |
| 40 | 31 | F | 19 | 2 | S | P | 6 | 6 | 6 | 6 | T | $\Psi$ | 0 | － | 2 | 2 | － | － | 3 |


|  |  | $\frac{8}{8}$ | ${ }_{4}^{4}$ |  | $\begin{aligned} & 4 \\ & 8 \\ & 8 \\ & 8 \\ & 8 \end{aligned}$ |  |  |  | Ira. Ji ano Study |  |  |  |  |  | Ac <br> 亳 葛 | ieve nent <br>  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 42 | 18 | 1 | 19 | 1. | T | $\nabla$ | 4 | 4 | 2 | － | I | $\underline{T}$ | － | － | － | － | 2 | 1 | 3 |
| 42 | 19 | F | 18 | 1 | A | P | 2 | 5 | － | 5 | $\mathbf{Y}$ | $\pm$ | － | － | － | － | 2 | 2 | 1 |
| 43 | 34 | F | 19 | 7 | 5 | B | 3 | 9 | 1 | $\frac{7}{3}$ | \％ | H | 0 | － | － | － | 1 | 4 | － |
| 44 | 49 | F | 20 | 3 | A | 篤 | 7 | 9 | 3 | 3 | Y | N | － | 0 | － | － | 2 | 2 | 3 |
| 45 | 11. | H | 18 | 1 | T | $B$ | 2 | 4 | 1 | 5 | I | Y | － | － | － | － | 2 | 1 | 2 |
| 45 | 25 | F | 19 | 1 | A | P | 5 | 8 | 5 | 8 | I | Y | 0 | － | － | － | 2 | 1 | 3 |
| 47 | 12 | 環 | 20 | 3 | B | $B$ | 9 | 9 | 1 | 1 | N | H | 0 | 0 | 1 | 2 | 1 | 1 | 3 |
| 48 | 16 | F | 19. | 1. | 5 | $\checkmark$ | 1 | 5 | $\pm$ | 2 | T | H | － | － | － | － | 3 | 1 | 3 |
| 49 | 40 | \％ | 23 | 1 | V | $p$ | 10 | － | 10 | － | $\underline{T}$ | 7 | － | － | － | － | 2 | 2 |  |
| 50 | 40 | F | 19 | 1 | A | P | 5 | 10 | 5 | 10 | $\pm$ | I | 0 | － | － | － | 1 | 1 | 2 |
| 51 | 20 | 7 | 20 | 1 | A | $p$ | 15 | 15 | 15 | 15 | $\pm$ | Y | 0 | － | － | $\square$ | 2 | 2 | 1 |
| 52 | $45^{\prime}$ | F | 18 | 1 | A | P | 12 | 12 | 12 | 12 | I | T | 0 | 4 | － | － | 1 | － | 3 |

TABLE 52

DATA FOR BUSTCLAMS TH SCHOOL C

|  | 管 | \％ |  |  |  |  |  | Apots ourcte sun |  | $\begin{aligned} & \text { 黄 } \\ & \text { 窝 } \\ & \stackrel{9}{4} \end{aligned}$ |  |  |  | $\begin{array}{r} \text { Acht } \\ \text { me } \\ \hline \end{array}$ | evo－ <br> nt <br>  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 133 | F | 20 | 3 | A | 里 | 7 | 7 | 2 | 2 | 7 | I | 0 | 0 | 2 | 2 | 1 | 3 |
| 257 | 迷 | 22 | 3 | T | $\nabla$ | 1 | 4 | 1 | 1 | x | I | 0 | － | 2 | 2 | 1 | 2 |
| 324 | H | 21 | 3 | B | $p$ | 7 | 7 | 7 | 7 | ＋ | I | 咅 | － | 2 | 2 | 2 | 2 |
| 434 | F | 27 | 3 | A | $p$ | $2 \frac{1}{2}$ | 5 | $2 \frac{2}{2}$ | 5 | H | I | 0 | － | 1 | 3 | 1 | 3 |
| 530 | M | 26 | 3 | T | 5 | 10 | 10 | 2 | 2 | I | N | 0 | 0 | 2 | 2 | 1 | 3 |
| 658 | H | 20 | 3 | B | 0 | 7 | 7 | 15 | 15 | I | I | 2 | 0 | 2 | 2 | 2 | 2 |
| 793 | F | 21 | 3 | 5 | 0 | 6 | 6 | 10 | 10 | 7 | $\underline{7}$ | 0 | － | 2 | 1 | 2 | 4 |
| 8.23 | 退 | 21 | 3 | T | 7 | 6. | 13 | 1. | 1 | 7 | $\Psi$ | － | 1 | 2 | 2 | 2 | 2 |
| 9102 | \％ | 19 | 3 | B | p | 6 | 12 | 6 | 12 | I | I | － | － | 1 | 1 | 2 | 1 |
| $10 \quad 24$ | F | 25 | 3 | 5 | \％ | 6 | 12 | 7 | 7 | I | 2 | 3 | 0 | 1 | 1 | 1 | 3 |

## table 52 （conciuded）

| $\begin{gathered} \text { 合 } \\ \frac{0}{0} \\ \text { on } \end{gathered}$ |  | 䈓 | $\underset{4}{9}$ |  | $\begin{aligned} & \text { 曹 } \\ & 8 \\ & 8 \\ & 8 \\ & \hline 8 \end{aligned}$ |  | Hde Apaqs - Sax |  | Kpmos cuetd sxt | Xri. Piano Bxpex. |  |  |  |  | Acht <br> men | eve－ nt |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 11. | 76 | F | 20 | 3 | A | P | 17 | $\underline{\square}$ | 17 | － | I | I | － | $\cdots$ | 2 | 2 | 1 | 3 |
| 12 | 91. | F | 20 | 3 | 5 | V | 2 | 6 | 35 | 15 | I | $\underline{1}$ | 0 | － | 1 | 1 | 2 | 4 |
| 13 | 42 | $T$ | 22 | 3 | 5 | P | 12 | 12 | 12 | 12 | I | T | $\underline{0}$ | － | 1 | 2 | 2 | 3 |
| 14 | 40 | \％ | 20 | 3 | 8 | \％ | 14 | $\cdots$ | 14 | $\square$ | Y | Y | 0 | － | 1 | 2 | 2 | 3 |
| 15 | 36 | 蓝 | 21 | 3 | T | 7 | 7 | 15 | 2 | 4 | H | I | 0 | － | 2 | 2 | 1 | 3 |
| 16 | 38 | F | 20 | 3 | $\because$ A | P | 0 | 0 | 8 | 8 | W | I | － | － | 2 | 2 | 1 | 3 |
| 17 | 78 | 1 | 20 | 3 | A | $V$ | 3 | 8 | 5 | 13 | I | V | 0 | － | 2 | 1 | 1 | 3 |
| 18 | 53 | F | 21. | 3 | S | V | 4 | 7 | 7 | 7 | I | I | $\cdots$ | $\square$ | 1. | 1 | 1 | 3 |
| 19 | 34 | F | 21 | 3 | $\mathbf{S}$ | $V$ | 7 | 8 | 2 | $-$ | 等 | H | 0 | － | 2 | 2 | 4 | 1 |
| 20 | 12 | 置 | 24 | 3 | O | B | 6 | 11. | 2 | 2 | 家 | 哏 | 0 | 0 | 2 | 2 | 1 | 3 |
| 21 | 38 | 榣 | 22 | 3 | B | 3 | 4 | 4 | － | 0 | 翟 | W | － | 1. | 2 | 3 | 1. | 4 |
| 22 | 37 | 翟 | 22 | 3 | － | B | 1 | 2 | $\cdots$ | 0 | 駺 | 要 | 0 | 1 | 2 | 2 | 4 | 4 |
| 23 | 37 | 咾 | 29. | 3 | B | 8 | 14 | 74 | 2 | 2 | 缼 | \＃ | 0 | $\frac{1}{2}$ | 2 | 2 | 2 | 2 |
| 24 | 46 | F | 20 | 3 | 3 | $p$ | 13 | 13 | 13 | 13 | I | I | 0 | $\underline{\sim}$ | 1 | 1 | 2 | 1 |
| 25 | 97 | F | 19 | 3 | A． | $p$ | 9 | 17 | 9 | 17 | I | T | $\longrightarrow$ | $\cdots$ | 2 | 1 | 2 | 3 |

## TABLE 53

DATA FOR MUSICIAWS IT SCHOOE D

|  | $\begin{aligned} & \text { 合 } \\ & \text {. } \\ & \text { 諸 } \\ & \text { E } \end{aligned}$ | 落 | 4． |  |  |  |  |  | $\text { Splas oxeTd } 8 x X$ |  | $\begin{aligned} & \text { 膏 } \\ & \text { 膏 } \\ & \text { 畐 } \end{aligned}$ |  |  |  | Ach m <br>  | ieve <br> ont <br>  <br> 雷 |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 1 | 33 | F | 19 | 2 | A | 5 | 11 | 17 | 1 | 1 | Y | 1 | 0 | － | 2 | 1 | 1 | 2 |
| 2 | 18. | F | 18 | 1 | 5 | V | 3 | 6 | 1 | 1. | \＃ | ， | － | － | 2 | 2 | 1 | 2 |
| 3 | 14. | F | 19 | 1 | A | B | 3 | 3 | 10 | 10 | 7 | $\pm$ | － | － | 1 | 1 | 1 | 2 |
| 4 | 30 | F | 18 | 1 | A | 0 | 4 | 4 | 7 | 7 | 7 | Y | － | － | 2 | 2 | 1 | 1. |
| 5 | 16 | F | 20 | 2 | － | $P$ | 10 | 10 | 10 | 10 | Y | Y | － | － | 1 | 2 | 1 | 2 |
| 6 | 13 | F | 18 | 1 | S | P | 9 | － | 9 | － | $\underline{Y}$ | － | － | － | 2 | 2 | 1 | 3 |
| 7 | 32 | F | 19 | 1 | － | P | 9 | 9 | 9 | 9 | Y | I | 5 | － | 2 | 1 | 1 | 1 |
| 8 | 25 | F | 17 | 1 | $s$ | V | 3 | 4 | 1 | 16 | ${ }^{W}$ | N | $\frac{7}{2}$ | － | 2 | 2 | 1 | 4 |
| 9 | 32 | F | 20 | 1 | A | P | 16 | 16 | 16 | 16. | I | I | － | － | 3 | 1 | 1 | 1 |
| 10 | 28 | F | 19 | 2 | A | F | 8 | 8 | 8 | 8 | I | $\Psi$ | － | － | 2 | 2 | 1 | 3 |

TABLE 53 （continued）

| $\begin{aligned} & \stackrel{3}{0} \\ & \stackrel{\rightharpoonup}{0} \\ & \stackrel{0}{\circ} \end{aligned}$ |  | \％ | ${ }_{5}^{8}$ |  |  |  | $\begin{aligned} & \text { 㯰 } \\ & \text { 营 } \\ & \text { 品 } \\ & \text { 咅 } \end{aligned}$ | $\begin{aligned} & \text { 喆 } \\ & \text { 娄 } \\ & \text { 部 } \\ & \frac{6}{4} \end{aligned}$ | $\begin{aligned} & H \\ & \frac{H}{6} \\ & \frac{9}{4} \\ & \frac{4}{4} \\ & \frac{4}{4} \end{aligned}$ |  |  |  |  |  | Ach： me <br> 总 瑶 | eve－ <br> nt <br> 品 喿 4 4 4 |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 11. | 25 | T | 19 | 2 | － | P | 7 | 7 | 7 | 7 | $\pm$ | I | $\cdots$ | － | 2 | 2 | 1 | 3 |
| 12 | 24 | F | 18 | 1 | S | P | 8 | － | 8 | － | $Y$ | Y | － | － | 3 | 2 | 4 | 3 |
| 13 | 16 | F | 18 | 1 | 5 | V | 3 | 5 | 4 | － | 4 | T | 0 | － | 2 | 3 | 1 | 3 |
| 14 | 24 | F | 20 | 2 | A | P | 8 | 8 | 8 | 8 | I | $\pm$ | － | － | 1 | 2 | 1 | 2 |
| 15 | 15 | F | － | 1 | A | 7 | 5 | － | 5 | － | $\pm$ | $\Psi$ | － | － | 2 | 3 | 4 | 2 |
| 16 | 72 | \％ | 18 | 2 | A | P | 8 | 8 | 8 | 3 | 7 | － | － | － | 2 |  | 2 |  |
| 17 | 6 | F | 19 | 1 | 5 | P | 1 | 1 | 1 | 1 | N | － | － | － | 2 | 2 | 1 | 3 |
| 18 | 25 | F | 19 | 2 | 5 | V | 1 | 2 | 3. | 3 | Y | \％ | $\cdots$ | － | 2 | 2 | 2 | 2 |
| 19 | 28 | F | 19 | 1 | 5 | $F$ | 9 | 9 | 9 | 9 | F | W | 0 | － | 1 | 2 | 1 | 4 |
| 20 | 26 | T | 18 | 1 | S | 5 | 9 | 9 | 4 |  | Y | I | － | 0 | 1. | 1 | 1 | － |
| 21 | 38 | T | 19 | 2 | S | $s$ | 8 | 8 | 8 | 8 | 7 | I | － | 0 | 2 | 2 | 2 | 3 |
| 22 | 53 | F | 19 | 2 | A． | V | 2 | 6 | － | － | \＃ | I | － | － | 1 | 1 | 2 | 2 |
| 23 | 59 | F | 17 | 1 | S | P | 12 | 12 | 12 | 12 | \％ | 7 | － | － | 2 | 1. | 2 | 4 |
| 24 | 29 | F | 19 | 2 | 5 | \％ | 11. | 11 | 5 | － | I | I | － | 0 | 2 | 2 | 1 | 3 |
| 25 | 28 | F | 19 | 2 | $s$ | 5 | 2 | 2 | 3 | 11. | T | I | 0 | 0 | 2 | 2 | 1 | 2 |

TABLE 53 (concluded)

| $\begin{aligned} & \text { B } \\ & \stackrel{8}{7} \\ & \text {. } \end{aligned}$ |  | 苞 | $\stackrel{8}{8}$ |  | $\begin{aligned} & 4 \\ & \stackrel{4}{4} \\ & 8 \\ & 8 \\ & 8 \end{aligned}$ |  | Yrs. Study PRU |  | Tres. Mano Stucy |  |  |  |  |  | Ach <br> me <br> 莒 | eve- <br> nt <br> 4 4 4 4 4 |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 26 | 20. | F | 18 | 2 | 5 | $P$ | 12 | 12 | 12 | 12 | T | T | - | - | 2 | 2 | 4 | 3 |
| 27 | 31. | F | 19 | 2 | 5 | $\nabla$ | 2 | 2 | - | - | I | Y | - | - | 2 | 1 | 1 | 2 |
| 28. | 32 | T | 19 | 1 | - | P | 11. | 14 | 14 | 14 | T | I | - | - | 1 | 1 | 1. | 2 |
| 29 | 27 | T | 19 | 2 | 5 | $\nabla$ | 3 | 5 | - | 6 | Y | I | - | - | 2 | 2 | 4 | 3 |
| 30 | 15 | F | 19 | 1 | 5 | V | 5 | - | 10 | - | $\underline{1}$ | (1) | - | - |  | 3 | . | 1 |
| 31. | 20 | F | 29 | 1 | A | 0 | 1 | 1 | 9 | - | I | 7 | 0 | $\cdots$ | 1 | I | - | 2 |
| 32 | 27 | $F$ | 19 | 2 | A | 5 | 10 | 15 | 1 | 3 | I | I | - | 2 | 2 | 3 | - | 3 |

## TABLE 54

DATA FOR GUSTCIARS IN SCHOOL $H$

| $\begin{aligned} & 8 \\ & 0 \\ & \text { 曷 } \\ & 0 \end{aligned}$ |  |  | 808 |  |  |  |  |  |  |  | $\begin{aligned} & \text { 德 } \\ & \text { 雷 } \\ & \text { 学 } \end{aligned}$ |  |  |  |  | ve－ nt <br>  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 1 | 19 | F | 19 | 1 | S | 5 | 6 | 6 | 2 | 2 | Y | － | 0 | － | － | － | 1 | 3 |
| 2 | 21 | 艮 | 23 | 1 | T | W | 5 | 5 | 3 | 3 | － | 1 | 3 | － | $-$ | － | 4 | 4 |
| 3 | 42 | M | 18. | 1 | T | B | 8 | 8 | 妾 | － | N | $\Psi$ | － | － | 1 | 1 | 2 | 3 |
| 4 | 32 | 楯 | 18 | 1 | T | B | 8 | 8 | 1 | 1. | \％ | I | 0 | 0 | 2 | 2 | 1. | 3 |
| 5 | 20 | F | 19 | 1 | S | $\nabla$ | 7 | 7 | 3 | 3 | $\pm$ | 7 | 0 | － | － | － | － | 3 |
| 6 | 17 | 塈 | 19 | 1 | V | B | 2 | 5 |  |  |  | $\pm$ | 0 | 0 | 2 |  |  |  |
| 7 | 29 | 唯 | 19 | 1 | B | B | 1 | 8 | 咅 | 咅 | U | I | － | － | 2 | 1 | 1 | 3 |
| 8 | 25 | 14 | 19 | 1 | B | B | 1 | 9 | $\frac{7}{7}$ | 1 | H | I | － | － | 2 | 1 | 1 | 3 |
| 9 | 19 | ［ | 19 | 1 | B | B | 1 | 9 | $\frac{2}{2}$ | － | W | $\underline{I}$ | 0 | 0 | 2 | 2 | 1 | 3 |
| 10 | 35 | \％ | 19 | 1 | B | 5 | 2 | 2 | 1 | 1 | N | I | 0 | 0 | 1 | 1 | 1 | 3 |

TABLB 54 (concluded)

mabie 55
data for mosicians in school F

|  | $\begin{aligned} & 0 \\ & 8 \\ & 8 \\ & 0 \\ & \mathbf{V}_{2} \\ & 2 \end{aligned}$ | \％ 6 | \％ |  | $\begin{aligned} & \overrightarrow{4} \\ & \text { N } \\ & \frac{1}{0} \\ & \frac{1}{8} \end{aligned}$ |  |  | 8 <br> 4 8 6 <br> म | Irs．Plano Study |  |  | $\begin{aligned} & \text { o } \\ & \text { H } \\ & \text { 最 } \\ & \end{aligned}$ |  |  | Ac <br> 免 怠 |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 1 | 15 | F | 22 | 3 | A | 5 | 3 | 4 | 5 | 5 | $\pm$ | Y | $\frac{1}{2}$ | 誊 | 1 | 1 | 1 | 3 |
| 2 | 27 | H | 30 | 3 | B | P | 2 | － | 2 | － | I | Y | \％ | － | 2 | 3 | 2 | 3 |
| 3 | 26 | ${ }^{\text {r }}$ | 28 | 3 | T | $B$ | 6 | 12 | 1 | 1 | W | 柯 | 1 | 2 | 2 | 2 | 1 | 3 |
| 4 | 76 | 等 | 28 | 3 | $B$ | $\nabla$ | 5 | 10 | 4 | 4 | T | \＃ | 7 | － | 1 | 1 | 2 | 2 |
| 5 | 48 | F | 22 | 3 | A． | 5 | 2 | 3 | 2 | 4 | Y | N | 0 | 0 | 1 | 1 | 1 | 4 |
| 6 | 25 | H | 25 | 3 | B | V | 7 | 10 | 2 | 4 | T | H | 0 | － | 1 | 1 | 1 | 3 |
| 7 | 82 | F | － | 3 | A | P | 19 | 19 | 19 | 19 | I | I | 1 | － | 1 | I | 2 | 3 |
| 8 | 40 | F | 25 | 3 | S | P | 12 | 15 | 12 | 15 | $\pm$ | － | 4 | － | 1 | 1 | 1 | 3 |
| 9 | 34 | M | 25 | 3 | T | B | 5 | 14 | $-$ | － | \％ | I | 0 | 2 | 3 | 1 | 1 | 2 |
| 10 | 31 | H | 23 | 3 | T | B | 1. | 2 | 3 | － | 3 | \％ | 1 | 1 | 1 | － | 1 | 3 |
| 11. | 27 | 㛵 | 42 | 3 | B | V | 15 | 25 | 2 | $\square$ | $\pm$ | － | － | － | 1 | 2 | 1 | 3 |

## TABLE 56

DATA FOR KUSTCIANS IN SCHOOL G

| $\begin{aligned} & \mathrm{B} \\ & \stackrel{y}{\circ} \\ & \text { 另 } \\ & \hline \end{aligned}$ | $\begin{aligned} & 8 \\ & 0 \\ & 0 \\ & \text { en } \\ & \text { 苞 } \\ & \text { E } \\ & \hline \end{aligned}$ | $\begin{aligned} & 8 \\ & 80 \end{aligned}$ | \％ |  |  |  |  |  |  |  | $\begin{aligned} & \text { 总 } \\ & \text { 品 } \\ & \text { \% } \end{aligned}$ |  |  |  | Ach <br> \％ \％ 8 |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 1 | 30 | T | 20 | 1 | 5 | $\nabla$ | 10 | 10 | 6 |  | I | H | 0 | － | 1 | 2 | 1 |
| 2 | 75 | F | 21. | 2 | 5 | $V$ | 2 | 7 | 2表 | 7 | I | Y | － | － | 1 | 1 | 2 |
| 3 | 24 | \％ | 20 | 1 | T | V | 1 | 1 | 2 | － | － | － | － | － | 1 | － | 3 |
| 4 | 12 | F | 19 | 1 | 5 | $\nabla$ | 1 | 6 | 10 | 10 | Y | I | 0 | － | 1 | 1 |  |
| 5 | 25 | F | 20 | 2 | S | V | 3 | 7 | 5 | 7 | $Y$ | 1 | － | － | 1 | 4 | 3 |
|  | 27 | F | 22 | 1 | $s$ | V | 4 | 5 | 5 | 5 | I | H | 1 | － | 1 | 4 | 3 |
| 7 | 66 | F | 48 | 3 | S | P | 8 | 43 | 8 | 43 | I | T | － | － | 2 |  |  |
| 8 | 31 | F | 20 | 2 | S | P | 13 | IH | 13 | 14 | Y | 7 | － | － | 1 | 2 | 3 |
| 9 | 27 | 3 | 23 | 2 | T | $\checkmark$ | 6 | 8 | 4 | 6 | I | H | － | － | 3 | 1 | 2 |
| 10 | 26 | F | 20 | 1 | 5 | V | 6 | 0 | 6 | 6 | Y | I | － | － | 1 | 1 | 3 |
| 11 | 29 | F | 21 | 2 | S | V | 6 | 9 | 5 | － | T | 17 | $\pm$ | － | 1 | 4 | － |
| 12 | 33 | F | 20 | 2 | S | $\nabla$ | 1 | 3 | 10 | － | I | T | 素 | － | 1 | 1 | 3 |
| 13 | 26 | F | 18 | 1 | S | V | 4 | 5 |  |  | I | － | 0 | － | 1 | 2 | 3 |
| 14 | 35 | F | 21 | 1 | A | $\nabla$ | 2 | 5 | $2 \frac{1}{2}$ | 27 | I | X | 3 | － | 2 | 1 | 2 |

TABLE 57

DATA FOR RUSTCIATS IN SGHOOL H


## TABLE 58

DATA FOR MUSICIAKS IT GRITERTON GROUP

| Subject | Test Score |
| :---: | :---: |
| 1 | 104 |
| 2 | 98 |
| 3 | 95 |
| 4 | 95 |
| 5 | 93 |
| 6 | 89 |
| 7 | 83 |
| 8 | 83 |
| 9 | 82 |
| 10 | 81 |
| 11 | 74 |
| 12 | 71. |
| 13 | 70 |
| 14 | 69 |
| 15 | 69 |
|  | 65 |
| 17 | 64 |
| 18 | 63 |
| 19 | 61 |
| 20 | 58 |
| 21. | 58 56 |
| 23 | 54 |
| 24 | 51 |
| 25 | 48 |

APPWDIX B
tixploratory Test

This is a test of your ability to correlate the tones and hamonies of choral music, as heard in performance, with their notation in the score. This ability is commonly exera cised during a reheareal situation in eliminating errors and securling the correct performance of the notes and chords.

Each musical passage consists of a phrase or more selectod from readily available choral music, such as might be used by the average choral group. There are no errors in the printed scores.

You will hear these passages performed, in a plano medlum, by means of a recording. In the performance of each excerpt one or more errors may be introduced, such as might be caused by inaccurate sight readers. After the number of the excerpt is announced, there will be an interval of approximately five seconds in which to study the score, after which the performance of the passage will be heard. As it is performed you are to listen for errors (harmonies or tones which do not agree with their notation in the score) and check each such error in the blank placed above each chord or note.

After an interval of approximately three seconds, the excerpt will be performed again, exactly like the first time. During this hearing you are to circle the head of the exact note or notes which have not been performed correctly. This second performance of the passage will be followed by an interval of approximately five seconds before the number of the next passage is announced.

In some instances, more than one error may occur in a chord. In such cases, after checking the blank, circie all the notes which have been performed incorrectly.

In the case of a chord performed incorrectily in which the apecific note in error is held into the following chord or chords, making them also incorrect, only the first wrongIy performed chord and note need be checked and circled.

It is an error on your part to check chords or circle notes which are performed correctiy.

This test is concerned only with the detection of discrepancies between the notation and the performance of notes and chords. The following are not to be considered as errors:

Tempo, rhythm, or phrasing.
Small alscrepancies in pitch, less than a half step.
Indistinct tones due to balence, blend, or recording defects.











ロロロロロ







白吕台台口 ロロ ロロロロ 名 名号


Mu－sic，when soft var－ces die，Vi－brates in the menn－o－ry，


Mu－sic，when soft roi－ces die，Vi－brates in the mavn－o－ry，






N. 51


№. 52


songs can wolease Thee, thronght thy bless-ed Son.




## APPSMDXX 0

## Frelininary Test, Fart 1. Reliability Data and Item Analystis Data

Flease answer the following questions

1. Age: $\qquad$
2. jex (check one): N
3. Bchool level (encircle the hiohest completed):

Hich school: grade 9 IO Il 12
Colleme: year 1 ? 34
Graduate: year 123
Are you in school now? Yes _ No __
4. Music theory completed (check hichest level onlyj:
:udiments
lifeh school harmony
College: flrst year harmony and ear trajnint
College: Second year harmony and ear trainine -
Check each of the follovine if comvleted:
Form and/or analysis. $\qquad$
Counterpoint $\qquad$
Conductine
Crchestration or instrumental arrancinc. $\qquad$
Choral arranfing $\qquad$
Composjticn $\qquad$
5. Musical experience:

Do you play the piano? Yes No
liow many years of formal piano stuतy":
How many years of pizno playine experience: $\qquad$
Vocal experience:
Uhat voice part do you sinp?(encircle one): S 1 I 3 Number of years formal voice study' Number of years singing experience, includine choral: Number of years choral directing experience? $\qquad$
Instrumental experience:
That is your principal instrument, if not piano or vojce"
liow many years of formal study on this?
Number of years playinc experience?
Number of years instrumental directing experience? $\qquad$

This is a teat of your ability to correlate the tones and harmonies of choral music, as heard in performance, with their notation in the score. This ability is commonly exercised by the music director during a rehearsal as he seeks to eliminate tonal errors and thus to secure the correct per-formance of the tones and harmonies as indicated in the score。

Each musical pasaage consists of a phrase or more selecta ed from readily available music, such as might be used by the average choral group. There are no errors in the printed scores.

You wlll hear these passages performed by means of a recording. In the performance of each excerpt one or more errors may be introduced, such as might be oaused by inaocurate sight reading by the singers. After the number of the excerpt is announced, the pitch of the beginning note or chord will be sounded on a pltch $p 1 \mathrm{pe}_{\mathrm{g}}$ followed by an interval of approximately ten seconds in winich to inspect the score. The performance will then be heard. As it is performed you are to 11 sten for errors (tonea or harmonies which do not agree with their printed notation), and check each such error in the blank placed above asch chord or note.

After an interval of approximately five seconds the excerpt will be performed again, exactly like the first time. With this hearing you are to circle the head of the speciric note or notes which have not been performed correctly. This second performance of the passage will be followed by an ine terval of approximately ilve geconds before the number of the next passage is announced.

In some instances, more than one note in a chord may be performed incorrectly. In such cases, after checking the blank above the chord, circle all the notes which have been performed incorrectly.

In the case of a chord performed incorrectly in which the apecific note in error is held into the following chord or chords, making them also incorrect, only the firgt wrongly performed chord and note need be checked and circled.

It is an error on your part to cheok chorde or alrele notes which are performea correctly.

The following are not to be considered as errors: Tempo, rhythim, or phrasing. Smali discrepencies in pitch, lesa than a half step. Indigtinct tones due to balance: bland, or recording derecta.








## $256$









$\mathrm{N}_{2} 44$

No. 45
Naw rest be- ncath nights shad - ows Main, beast, tow,




## TABLE 59

PRELLHIMARY TEST, PART I: SCORES OBTATNED FOR SFIIT HALUSS OF WEST, USIM AITTENATE ITEMS, FOR USE IM ESTIMATIEG TEST RELIABIIITY

| $\frac{0 d d}{X}$ | Even $\Psi$ | $\begin{gathered} \text { Odd } \\ \mathrm{X} \end{gathered}$ | Even I |
| :---: | :---: | :---: | :---: |
| 44 | 41 | 63 | 62 |
| 56 | 50 | 24 | 28 |
| 15 | 18 | 25 | 37 |
| 45 | 35 | 18 | 22 |
| 31 | 35 | 11. | 12 |
| 16 | 14 | 19 | 9 |
| 15 | 14 | 24 | 23 |
| 21. | 11 | 17 | 24 |
| 23 | 15 | 19 | 13 |
| 14 | 13 | 7 | 7 |
| 7 | 5 | 11 | 12 |
| 11. | 10 | 20 | 17 |
| 12 | 11. | 36 | 32 |
| 53 | 55 | 39 | 31 |
| 27 | 24 | 55 | 50 |
| 46 | 29 | 53 | 47 |
| 20 | 25 |  |  |
| 18 | 12 |  |  |
| 21. | 25 | $\Sigma \mathrm{X}=970$ | $\Sigma Y 900$ |
| 34 | 32 |  |  |
|  |  |  |  |

## TABLS 60

TMDICES OR DTFFICULTY ANO DISGRTMTHATMG PORER OF INDIVIDTAL ITEXS: PRLIMINARY TEST, PART I

| Excerpt | Item | Percent Passing |  |  | Phi | Items discriminating at given levels of confíidence ${ }^{1}$ |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  |  | Group | Hale | Lower <br> Holf |  | $\frac{\text { given } 1 e}{.01}$ |  | $\frac{1 \text { Hedex }}{.10}$ |
| $\frac{1}{2}$ | 1 | 47.2 | 55.5 | 38.8 | 170 |  |  |  |
|  | 1 | 0.3 | 11.1 | 5.5 | 090 |  |  |  |
|  | 2 | 22.2 | 38.8 | 5.5 | 395 |  | $x$ |  |
|  | 3 | 8.3 | 16.6 | 0.0 | 305 |  |  | $\mathbf{x}$ |
|  | 4 | 33.3 | 44.4 | 22.2 | 234 |  |  |  |
|  | 5 | 27.7 | 44.4 | 11.1 | 370 |  | $x$ |  |
| 3 | 1 | 5.5 | 11.2 | 0.0 | 241 |  |  |  |
|  | 2 | 13.8 | 27.7 | 0.0 | 403 |  | $x$ |  |
|  | 3 | 19.4 | 33.3 | 5.5 | 341 |  | $x$ |  |
| 4 | 1 | 25.0 | 38.8 | 11.1 | 323 |  |  | x |
|  | 2 | 66.6 | 88.8 | 44.4 | 477 | $x$ |  |  |
| 5 | 1 | 27.7 | 50.0 | 5.5 | 490 | x |  |  |
|  | 2 | 61.1 | 88.8 | 33.3 | 574 | \% |  |  |
| 6 | 1 | 30.5 | 44.4 | 16.6 | 293 |  |  | x |
|  | 2 | 27.7 | 38.8 | 16.6 | 245 |  |  |  |
| 7 | 1 | 38.8 | 38.8 | 38.8 | 0 |  |  |  |
|  | 1 | 66.6 | 88.8 | 44.4 | 477 | $x$ |  |  |
|  | 2 | 38.8 | 55.5 | 22.2 | 349 |  | x |  |
|  | 3 | 13.8 | 22.2 | 5.5 | 231 |  |  |  |
|  | 4 | 5.5 | 31.1 | 0.0 | 24. |  |  |  |
| 9 | 1 | 30.5 | 38.8 | 22.2 | 185 |  |  |  |
|  | 2 | 50.0 | 61.1 | 38.8 | 220 |  |  |  |
| 10 | 1 | 75.0 | 88.8 | 61.1 | 323 |  |  | x |
| 11 | 1 | 66.6 | 88.8 | 4.4 | 177 | x |  |  |
| 12 | 1 | 44.4 | 72.2 | 16.6 | 553 | x |  |  |
|  | 2 | 47.2 | 83.3 | 11.1 | 721 | x |  |  |
|  | 3 | 38.8 | 72.2 | 5.5 | 677 | x |  |  |
|  | 4* | 58.3 | 50.0 | 66.6 | 173 |  |  |  |
|  | 5 | 50.0 | 83.3 | 16.6 | 660 | $x$ |  |  |

$1_{V a l u e s ~ o f ~ p h i ~ r e q u i r e d ~ f o r ~ i t e m s ~ t o ~ d i s c r i m i n a t e ~ a t ~ t h e ~}^{01,} .05$, and . 10 levels of confldence are . 429, . 327, and . 274 respectively.

拱tems so marked discriminated negatively.

TABLE " 60 (continuod)

| Excerpt | Item | Percent Passing |  |  | Ph | Items discrininating at given lovels of confidence |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  |  | Total Group | thpor | hower Half |  |  |  |  |
|  |  |  |  |  |  | . 01 | . 05 | .10 |
| $\begin{aligned} & 13 \\ & 14 \end{aligned}$ | 1 | 52.7 | 55.5 | 50.0 | 060 |  |  |  |
|  | 1 | 33.3 | 38.8 | 27.7 | 237. |  |  |  |
|  | 2 | 38.8 | 66.6 | 11.1 | 574 | $x$ |  |  |
| 15 | 1 | 58.3 | 72.2 | 44.4 | 284 |  |  | x |
|  | 2 | 38.8 | 72.2 | 5.5 | 677 | x |  |  |
| 16 | 1 | 66.6 | 77.8 | 55.5 | 234 |  |  |  |
|  | 2 | 11.1 | 16.6 | 5.5 | 172 |  |  |  |
|  | 3 | 5.5 | 5.0 | 5.0 | 0 |  |  |  |
|  | 4 | 5.5 | 5.0 | 5.0 | 0 |  |  |  |
| 17 | 1 | 30.5 | 50.0 | 11.1 | 424 |  | x |  |
|  | 2 | 22.2 | 38.8 | 5.5 | 395 |  | X |  |
|  | 3 | 50.0 | 50.0 | 50.0 | 0. |  |  |  |
| 18 | 1 | 19.4 | 22.2 | 16.6 | 063 |  |  |  |
| 19 | 1 | 22.2 | 27.7 | 26.6 | 132. |  |  |  |
|  | 2 | 13.8 | 22.2 | 5.5 | 231 |  |  |  |
|  | 3 | 19.4 | 38.8 | 0. | 492 | x |  |  |
|  | 4 | 72.2 | 94.4 | 50.0 | 490 | * |  |  |
| 20 | 1 | 4.4 | 55.5 | 33.3 | 231. |  |  |  |
| 21. | 1 | 44.4 | 77.7 | 11.1 | 674 | * |  |  |
| 22 | 1 | 42.6 | 61.1 | 22.2 | 396 |  | $x$ |  |
|  | 2 | 25.0 | 38.8 | 11.1 | 323 |  |  | $x$ |
|  | 3 | 30.5 | 33.3 | 27.7 | 054 |  |  |  |
|  | 4 | 33.3 | 4.4 | 22.2 | 234 |  |  |  |
| 23 | 1 | 63.8 | 66.6 | 61.1 | 063 |  |  |  |
| 24 | 1 | 47.2 | 77.7 | 16.6 | 611 | $x$ |  |  |
|  | 2 | 25.0 | 38.8 | 11.1 | 323 |  |  | X |
|  | 3 | 50.0 | 61.1 | 38.8 | 220 |  |  |  |
| 25 | 1 | 55.5 | 83.3 | 27.7 | 553 | $x$ |  |  |
|  | 2 | 31.1 | 16.5 | 5.5 | 172 |  |  |  |
|  | 3 | 27.7 | 4.4 | 11.1 | 370 |  | x |  |
|  | 4 | 27.7 | 50.0 | 5.5 | 490 | * |  |  |
|  | 5 | 33.3 | 55.5 | 11.1 | 477 | I |  |  |
|  | 6 | 27.7 | 414 | 11.1 | 370 |  | X |  |
| 27 | 1 | 36.1 | 61.1 | 11.1 | 521 | * |  |  |
|  | 1 | 30.5 | 50.0 | $\frac{11.1}{22.2}$ | 424 |  | $\mathbf{x}$ $\mathbf{x}$ |  |
|  | 2 | 41.6 | 61.1 | 22.2 | 396 |  | X |  |
|  | 3 | 44.4 | 55.5 | 33.3 | 231. |  |  |  |

TABLE 60 (continued)

| Excerpt | Item | Percent Passing |  |  | Phit | Itens diacriminating at given levels of confidence |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  |  | Total | Upper | Lower |  |  |  |  |
|  |  | Group | Half | Hald |  | . 01 | . 05 | .10 |
| 28 | 1 | 55.5 | 66.6 | 44.4 | 231 |  |  |  |
|  | 2 | 13.8 | 27.0 | 0.0 | 403 |  | x |  |
|  | 3 | 58.3 | 72.2 | 44.4 | 284 |  |  | * |
|  | 4 | 36.1 | 44.4 | 27.7 | 167 |  |  |  |
|  | 5 | 58.3 | 88.8 | 27.7 | 619 | x |  |  |
| 29 | 1. | 47.2 | 61.1 | 33.3 | 281 |  |  | $x$ |
|  | 2 | 38.8 | 61.1 | 16.6 | 451 | \% |  |  |
| 30 | 1 | 44.4 | 61.1 | 27.7 | 332 |  | $\mathbf{x}$ |  |
|  | 2* | 25.0 | 22.2 | 27.7 | 069 |  |  |  |
|  | 3 | 50.0 | 55.5 | 4.404 | 110 |  |  |  |
| 31. | 1 | 13.8 | 16.6 | 11.1 | 086 |  |  |  |
|  | 2 | 14.4 | 66.6 | 22.2 | 453 | $x$ |  |  |
| 32 | 1 | 50.0 | 61.1 | 38.9 | 220 |  |  |  |
|  | 2 | 58.3 | 88.8 | 27.7 | 619 | $x$ |  |  |
|  | 3 | 25.0 | 33.3 | 16.6 | 185 |  |  |  |
| 33 | 1 | 13.8 | 16.6 | 71.1 | 086 |  |  |  |
|  | 2 | 41.6 | 72.2 | 31.1 | 619 | x |  |  |
|  | 3 | 61.1 | 83.3 | 38.9 | 451 | $\mathbf{x}$ |  |  |
|  | 1 | 58.3 | 72.2 | 4.4 | 284 |  |  | X |
| $35^{\prime}$ | 1 | 69.4 | 94.4 | 44.4 | 512 | $x$ |  |  |
|  | 2 | 25.0 | 38.8 | 11.1 | 323 |  |  | x |
|  | 3 | 25.0 | 50.0 | 0.0 | 577 | x |  |  |
|  | 4* | 27.7 | 22.2 | 33.3 | 1.23 |  |  |  |
| 36 | 1 | 25.0 | 33.3 | 16.6 | 185 |  |  |  |
|  | 2 | 16.6 | 27.7 | 5.5 | 293 |  |  | $\boldsymbol{x}$ |
| 37 | 1 | 11.6 | 77.7 | 5.5 | 729 | $x$ |  |  |
|  | 2 | 38.8 | 72.2 | 5.5 | 677 | $x$ |  |  |
|  | 3 | 22.2 | 38.8 | 5.5 | 395 |  | X |  |
| 38 | 1 | 19.4 | 33.3 | 5.5 | 341 |  | x |  |
|  | 2 | 52.7 | 72.2 | 33.3 | 390 |  | * |  |
|  | 3* | 13.8 | 11.1 | 16.6 | 086 |  |  |  |
|  | 4 | 11.1 | 16.6 | 5.5 | 172 |  |  |  |
| 39 | 1 | 36.1 | 50.0 | 22.2 | 292 |  |  | $x$ |
|  | 2 | 33.3 | 38.8 | 27.7 | 117 |  |  |  |
| 40 | 1 | 30.5 | 61.1 | 0.0 | 662 | $x$ |  |  |
|  | 2 | 27.7 | 55.5 | 0.0 | 624 | $x$ |  |  |
|  | 3 | 16.6 | 33.3 | 0.0 | 445 | $x$ |  |  |
|  | 4 | 16.6 | 33.3 | 0.0 | 4 | * |  |  |
|  | 5 | 38.8 | 38.8 | 12.1 | 323 |  |  | x |

TABLE 60 (contrmiea)

| Excerpt | Item | Percent Passing |  |  | Phi | Items diseriminating at given levels of confidence |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  |  | Total Upper Lower |  |  |  |  |  |  |
|  |  | Group | Half | Half |  | . 01 | . 05 | .10 |
|  | 6 | 33.3 | 50.0 | 16.6 | 350 |  | * |  |
|  | 7 | 25.0 | L4.4 | 5.5 | 439 | $\mathbf{x}$ |  |  |
| 41 | 1 | 22.2 | 4.0 | 0.0 | 531 | x |  |  |
|  | 2 | 38.8 | 61.1 | 36.6 | 451 | $x$ |  |  |
|  | 3 | 80.5 | 100.0 | 61.1 | 492 | $\mathbf{x}$ |  |  |
|  | 4 | 47.2 | 55.5 | 38.8 | 170 |  |  |  |
| 42 | 1 | 13.8 | 22.2 | 5.5 | 231 |  |  |  |
|  | 2 | 11.1 | 22.2 | 0.0 | 352 |  | x |  |
|  | 3 | 16.6 | 33.3 | 0.0 | 14.5 | * |  |  |
|  | 4 | 13.8 | 27.7 | 0.0 | 403 |  | $x$ |  |
|  | 5 | 13.8 | 22.2 | 5.5 | 231 |  |  |  |
|  | 6 | 44.4 | 72.2 | 16.6 | 553 | $x$ |  |  |
| 43 | 1 | 38.8 | 61.1 | 16.6 | 451 | $\mathbf{x}$ |  |  |
|  | 2 | 33.3 | 50.0 | 16.6 | 350 |  | $x$ |  |
|  | 3 | 50.0 | 83.3 | 16.6 | 660 | $x$ |  |  |
|  | 4 | 55.5 | 77.7 | 33.3 | 453 | $\mathbf{x}$ |  |  |
| 44 | 1 | 47.2 | 77.7 | 16.6 | 671 | x |  |  |
|  | 2 | 47.2 | 55.5 | 38.8 | 170 |  |  |  |
|  | 3 | 52.7 | 83.3 | 22.2 | 611 | $x$ |  |  |
|  | 4 | 61.1 | 88.8 | 33.3 | 574 | $\mathbf{x}$ |  |  |
|  | 5 | 11.1 | 22.0 | 0.0 | 352 |  | $x$ |  |
|  | 6 | 11.1 | 16.6 | 5.5 | 172 |  |  |  |
|  | 7 | 66.6 | 83.3 | 50.0 | 350 |  | $x$ |  |
| 45 | 1 | 47.2 | 77.7 | 36.6 | 611 | x |  |  |
|  | 2 | 25.0 | 4.4 .4 | 5.5 | 439 | x |  |  |
|  | 3 | 66.6 | 72.2 | 61.1 | 137 |  |  |  |
|  | 4 | 38.8 | 44.4 | 33.3 | 113 |  |  |  |
|  | 5 | 4.4 | 72.2 | 16.6 | 553 | $x$ |  |  |
|  | 6 | 19.4 | 33.3 | 5.5 | 341 |  | x |  |
|  | 7 | 52.7 | 83.3 | 22.2 | 611 | x |  |  |
|  | 8 | 55.5 | 77.7 | 33.3 | 453 | x |  |  |
|  | 9 | 37.1 | 22.2 | 0.0 | 352 |  | $x$ |  |
|  | 10 | 19.4 | 38.8 | 0.0 | 492 | x |  |  |
|  | 11 | 22.2 | 38.8 | 5.5 | 395 |  | * |  |
|  | 12 | 25.0 | 50.0 | 0.0 | 577 | z |  |  |
|  | 13 | 52.7 | 83.3 | 22.2 | 611 | X |  |  |
| 46 | 1 | 25.0 | 4.4 | 5.5 | 439 | $x$ |  |  |
|  | 2 | 11.1 | 22.2 | 0.0 | 352 |  | $x$ |  |

## TABIE 60 (concluded)

| Excerpt | Item | Percent Passing |  |  | Phi | Items discriminating at given levels of confidence |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  |  | Total <br> Group | $\begin{aligned} & \text { Upper } \\ & \text { Hale } \end{aligned}$ | Lover <br> Half |  |  |  |  |
|  |  |  |  |  |  | . 01 | . 05 | .10 |
| 47 | 1 | 33.3 | 33.3 | 33.3 | 0 |  |  |  |
|  | 2 | 13.8 | 22.2 | 5.5 | 231 |  |  |  |
|  | 3 | 16.6 | 27.7 | 5.5 | 293 |  |  | 2 |
| 48 | 1 | 2.7 | 5.5 | 0 | 176 |  |  |  |
|  | 2 | 8.3 | 16.6 | 0 | 305 |  | $\pm$ |  |
|  | 3 | 22.2 | 44.4 | 0 | 531 | $x$ |  |  |
|  | 4 | 19.4 | 30.8 | 0 | 492 | \% |  |  |

## APPBNDIX D

## Preliminary Test, Part 2, Relifability Data, and Item Analysis Data

This is a test of your ability to correlate the tones and harmonies of choral music, as heard in performance, with their notation in the score. This ability is commonly exercised by the music director during a rehearsal as he seeks to eliminate tonal errors and thus to secure the correct performance of the tones and harmonies as indicated by the score.

Each musical passage consists of a phrase or more suto lected from readily available music, such as might be used by the average choral group. There are no errors in the printed scores.

You will hear these pasazges performed by means of a recording. In the performance of each excerpt one or more errors may be introduced, such as might be caused by inaccurate sight reading by the singerg. Following the announcement of the number of the excerpt, there will be an interval of approximately five seconds in which to inspect the score. The hummed pitch of the starting note or chord will then be heard, followed shortly by the first performance of the passage. After another interval of approximately five seconds, the starting pitch and the excerpt will be performed again, exactiy like the first time. After the second performance of the passage there will be an interval of approximately five seconds before the number of the next passage is announced.

During the performances of each musical excerpt you are to listen for errors in performance (tones or harmonies which do not agree with their printed notation), and check each such error in the blank placed above each chord or note。

In some instances, more than one note in a chord may be performed incorrectiy.

In the case of a chord performed incorrectly in which the specific note in error is sustained into the following chord or chords, making them also incorrect, only the blank above the first wrongly performed chord needs to be ohecked.

The following are not to be considered as errors: Tempo, rhythm, or phrasing. Small discrepancies in pitch, less than a half step. Indistinct tones due to balance, blend, or recording defects.

> 范.






No. 14

## TABLE 61

PRELTMTNARY TEST, PART 2y ScORES OBTATABD FOR SPLIT HALVES
 TEST RELJABILITY

| $\frac{0 d d}{x}$ | Eren I |
| :---: | :---: |
| 16 | 14 |
| 16 | 14 |
| 10 | 5 |
| 9 | 4 |
| 14 | 10 |
| 12 | 7 |
| 13 | 10 |
| 11 | 10 |
| 11 | 10 |
| 7 | 4 |
| 8 | 4 |
| 10 | 9 |
| 3 | 2 |
| 20 | 15 |
| 19 | 20 |
| 15 | 36 |
| 10 | 5 |
| 16 | $1{ }_{4}$ |
| 17 | 13 |
| IH | 9 |
| 12 | 10 |
| 12 | 13 |
| 7 | 5 |
| 8 | 10 |
| 9 | 8 |
| $\begin{array}{r}5 \\ 3 \\ 16 \\ \hline\end{array}$ | $\begin{array}{r}5 \\ 2 \\ 74 \\ \hline\end{array}$ |
| 工可 323 | $\Sigma \Psi=262$ |

IMDICES OF DHFIGULTY AND DTSGRTMMATIM POMER OF INDIVIDUAL ITENS: PGETMIMARY TEST, PART 2

| Excerpt | Item | Percent Passing |  |  | Phat | Items discriminating at Eiven levels of confidence ${ }^{1}$ |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  |  | Total | Upper | Lorer |  |  |  |  |
|  |  | Group | Half | Hale |  | . 11 | . 05 | . 10 |
| 1. | 1 | 68 | 100 | 36 | 686 | $\mathbf{x}$ |  |  |
|  | 2 | 54 | 64 | 43 | 211 |  |  |  |
|  | 3 | 36 | 36 | 36 | 0 |  |  |  |
| 2 | 1 | 71 | 100 | 43 | 631 | $\pm$ |  |  |
|  | 2 | 36 | 64 | 7 | 596 | x |  |  |
|  | 3 | 18 | 36 | 0 | 469 |  | $x$ |  |
| 3 | 1 | 64 | 79 | 50 | 303 |  |  |  |
|  | 1 | 57 | 79 | 36 | 435 |  | x |  |
|  | 2 | 32 | 36 | 29 | 075 |  |  |  |
|  | 3 | 50 | 64 | 36 | 280 |  |  |  |
|  | 4 | 50 | 79 | 21. | 50 | $x$ |  |  |
|  | 5 | 46 | 79 | 34. | 652 | $x$ |  |  |
|  | 6 | 76 | 86 | 71 | 183 |  |  |  |
| 5 | 1 | 57 | 86 | 29 | 577 | $x$ |  |  |
|  | 2 | 61 | 86 | 36 | 513 | $x$ |  |  |
|  | 3 | 32 | 50 | 14 | 386 |  | x |  |
|  | 4 | 50 | 79 | 21 | 580 | $x$ |  |  |
|  | 5 | 54 | 79 | 29 | 502 | $x$ |  |  |
| 6 | 1 | 75 | 93 | 57 | 416 |  | x |  |
| 7 | 1 | 60 | 71 | 64 | 075 |  |  |  |
|  | 2 | 27 | 21 | 21 | 0 |  |  |  |
|  | 3 | 36 | 57 | 14 | 449 |  | $x$ |  |
|  | 4 | 68 | 71 | 64 | 075 |  |  |  |
| 8 | 1* | 4.6 | 43 | 50 | 070 |  |  |  |
| 9 | 1 | 64 | 72 | 57 | 1.6 |  |  |  |
|  | 2 | 36 | 71 | 0 | 7 TH | 3 |  |  |
|  | 3 | 57 | 72 | 43 | 283 |  |  |  |
| 10 | 1 | 54 | 79 | 29 | 502 | $x$ |  |  |

IValues of phi required for items to discriminate at the $.01, .05$, and . 10 levels of confidence are . 487, 370 , and .310 respectively. Ninis item atscriminated negatively.

TABLS 62 (concluded)


## APPEWDIX E

# Final Test and Questionnaire, Item Selection Data, Analysis of 維配 Quality and Function of Test Items, and <br> Rellability Data 

1. Name
2. Sex: M F 3. Age $\qquad$
3. School level at present (oircle one):

College: year 123
Graduate school: year 123
(If not in school now, indicate by a oheck here , and oircle the highest
level which you have completed. How many years since you were in school? $\qquad$
5. Mark with a plus sign (4) the courses whioh you have completed, and with a check mark $(\sqrt{ })$ those which you are taking at present:

High school harmony
lst year college harmony and ear training
lst year keyboard harmony (if a separate ourse) $\qquad$
2nd year harmony and ear training
2nd year keyboard harmony (if a separate oourse)
Conducting Form and analysis Counterpoint _
Orchestration or instrumental arranging_ Choral arranging
composition $\qquad$ Other theory (specify and mark)
6. Piano experience:

Can you play hymn tunes and simple accompaniments at sight? Yes No.
Can you imprōise some type of piano accompaniment for simple melodies? Yes No
Is piano your principal instrument? Yes $\qquad$ No $\qquad$
How many years of formal piano study?
How many years of piano playing experience? $\qquad$
7. Vocal experience:

What voice part do you sing? (Circle one): S A T B
Is voice your principal performing medium? Yos No $\qquad$
How many years of formal voice study?
How many years of singing experience, including choral?
How many years of choral directing experienco?
8. Instrumental experience:

What is your principal instrument, if not piano or voice?
How many years of formal study on this?
How many years of playing experience? $\qquad$
How many years of instrumental direoting experience?
What other instruments do you play?
9. Please indioate your achiovement, aocording to your grades, in the last (most
rcoent) term in whioh you took the following coursos. (If you are a beginning student and have not yet completed a term, indicate what you believo your prosent aohievement to bo.)

Harmony: Above avoraga $\qquad$ Average / Below avorage
$\qquad$
Ear training: Above avcrage $\qquad$
$\qquad$ Bolow average $\qquad$
Sight singing (if a separatc oourse):
Above average Average $\qquad$ Bolow average $\qquad$
Kcyboard harmony (if a scparate course):
Above averago Avcrago Bolow average
All tho above (if togethor in a single course): -
Above average $\qquad$ Average $\qquad$ Below averago $\qquad$
In. Plcaso answer after the tost has beon administerod:
a. How did you liston for aocuraoy of harmonic offect? (Check onc):
(1) Listencd to onc or two parts ate a time
(2) Listencd to oomplate ohord quality (major, minor, oto.)
(3) Combination of (1) and (2). If you check this, which mothod did you use the moro: (1) (2)
(4) Listancd primaril and molodic flow
If you oheok this, what use did you make of
mathod (1): some uso little uso no usc mothod (2): somo uso_ littlo uso no usc -
b. To what extent did you analyze the ohords or recognize their function as being I, IV, V, eto.? (Chcok onc):

Most of the time
Frequently
Only oocasionally
Nover $\qquad$

This is a test of your ability to correlate the sounds of music, which you will hear performed, with their notation in the printed scores which you have before you. The exercise of this ability is a common experience of musicians as they seek to eliminate errors and secure the correct performance of the music.

This test is concerned only with the detection of melodic and harmonic differences between the notation of the music and its performance.

The test consists of thirty excerpts of choral music, each a phrase or more in length, selected from music such as the average choir might use. Each excerpt appears as the composer or arranger intended; there are no changes in notation. Any differences between score and performance are due to errors in performance alone.

Under each note or chord of the music appears a small square. As each selection is performed, melodic or harmonic errors may be introduced, such as might be caused by inaccurate sight reading on the part of the singers. Wherever you detect a note or chord which is not performed as it should be according to the notation in the score, you are to indicate this as an error by placing a check mark in the appropriate square.

Each musical selection will be performed twice, as follows: The number of the selection will be announced, followed by an interval of approximately five seconds in which to make a preliminary inspection of the score. The pitch of the note or chord on which the singers will start will then be hummed, followed shortly by the vocal performance. After an interval of approximately three seconds the pitch and music will be heard agein, exactly like the first time. There will then be another interval of approximately three seconds before the number of the next selection is announced.

There may be more than one error in the performance of each excerpt; there may be several, and in the longer ones there may be more. Selection number 17, for example, is a complete composition and may contain a number of errors.

More than one voice may be in error in the same chord.
It will be an error on your part to check the square under a note or chord which has been performed correctly.

If you change your mind about a check mark, either erase it or scratch through it several times with your pencil.

The following are not to be considered as errors:
(1) Variations of time, rhythm, or phrasing.
(2) Small pitch variations which are distinctly less than a halfstep. (But a note sung as a half-step which should have been a whole-step, for example, would be an error.)
-(0) Tunco wircir are maistinct aue to the balance or blend of the voices in the recorded performance.
(4) Occasional noises which are obviously due to mechanical defects in the recording and are unrelated to the music.

Turn now to the first page of the test, and listen to the example as you study the score. The check mark indicates that the chord directly above it was not performed correctly according to the notation. The test will proceed immediately after the example has been played.

Remember, any differences between score and Derformance are due to errors in performance alone; the music notation appears as the composer intended.

This test of tie ability to correlate the sound and notation of music is designed for misic anc music education majors. It is eppropriate for all levels of collece music classes, such as first and second year theory, conducting, counterpoint, form and analysis, arrangine, and composition. Please acminister the test to as many such classes as possible, even inclucinc those offered on the cracuate level, if available. The test may ke acministered individually.

Prior to administration of the test, the reel of tape should ke placed in tree machine to be used, and played at a speed setting of $7 \frac{1}{2}$ inches per second for a length of time sufficient for you to determine best adjustments for tone and volume. This pre-test playine of the tape will also serve to familiarize you with the nature of the test items.

Each person taking the test should have a pencil, and should ke scated at a desk or takle in order that the long forms may be hanuled easily.

As soon as the tests are distributed, the examinees should be instructed to inswer the questionnaire as completely and accurately as possible, except for question INo. 10. This is important, as the tost results cannot be evaluated without this information. Please emphasize that all indiviaual test results anci information will be confidential.

Question Ho. 9 calls for the examinee to check one of three categories of achievement, accordine to the marks he received, in the most recent term of the courses which he has completed. Grades of "A" or "B" or their equivalent are to be considered as "avove average", "C" or its equivalent as "everace", and "D" or "F" or their equivalent as "below average".

When each person has completed the questionnaire, all should turn to the pace of instructions and follovi silently as jou read the instructions aloud. Then ask if there are any questions. If there are, it vould be advisable to read the instructions again.

The playing of the tape may now begin.
When the last test item has been completed, please remind the examinees to answer question No. 10.

## SPECIAL REQUEST

It will ve greatly appreciated if you will inspect question iio. 9 on the questionnaire of each person whose prior achievement or quality of work is known to you, and verify the levels of achicvement wich have veen checked. It is necessary only to recall vil.ctior tie individual's work was above average, average, or below average. If you recall specific srades, please refer to the directions given above regaraine that question. If an individual kas checked what you know to ke a wrong category, ploase circle that mark and checli what shoulc have been his correct response. Please indicate the questionaires on which you have verified this iton, by placing a check mark in front of question No. 1 (the inuividual's name).


1


2


3


5









Bo－ne past－or pan－is ve－re Je－su nos－tri mi se
$\square \square \square \square \square_{2}$
$\square \square \frac{\square}{3}$
ㅁ
$\square \square \square \square$
$\square \square \square \square$
ロロロ ロ

Thou Son of God，the ho－14 Fa－ther，ho－14，heav＇n 14 One，Thou bles－sed Je－suchrist


22







My buckets andpails are shining and clean, for the finest milk you have ever seen



$\square \square \square \square \square \square \square \square \square \square \square$


TABLE 63

EXCERPTS AND ITETS SELECTED FOR USE TN FINAL FORM OF TEST, INGLUDING INDICES OF DISCRIUTNATING POWER AND LIEVEL OF DIFFTCULTY OF ITEMS

| Excerpt | Source: Prelim. Test and. Excerpt | Item | Level of Confim dence in DiscrinInation | Level of Difficulty (\% Total group passing) |
| :---: | :---: | :---: | :---: | :---: |
| 1 | 2-1 | 1 | . 01 | . 68 |
|  |  | 2 | . 01 | . 54 |
|  |  | 3 | - | . 36 |
| 2 | $2-2$ | 1 | . 01 | . 71 |
|  |  | 2 | . 01 | . 36 |
|  |  | 3 | . 05 | .18 |
| 3 |  | 1 | . 05 | . 75 |
|  | $2-5$ | 1 | . 01 | . 57 |
|  |  | 2 | . 01 | . 61 |
|  |  | 3 | . 05 | . 32 |
|  |  | 4 | . 01 | . 50 |
|  |  | 5 | . 01 | . 54 |
| 5 | $2-4$ | 1 | . 05 | . 57 |
|  |  | 2 | - | . 32 |
|  |  | 3 | - | . 50 |
|  |  | 4 | .01 | . 50 |
|  |  | 5 | . 01 | . 46 |
|  |  | 6 | - | . 79 |
| 6 | 1-11 | 1 | . 01 | . 67 |
| 7 | 2-9 | 1 | - | .64 |
|  |  | 2 | .01 | . 36 |
|  |  | 3 | - | .57 |
| 8 | 1-19 | 1 | - | . 22 |
|  |  | 2 | - | .14 |
|  |  | 3 | . 01 | . 19 |
|  |  | 4 | . 01 | . 72 |
| $\begin{array}{r} 9 \\ * 10 \end{array}$ | 1-21 | 1 | . 01 | . 44 |
|  | 1-33 | 1 | . 01 | . 42 |
|  |  | 2 | . 01 | . 61 |

WThls excerpt was revised and remrecorded to eliminate the original first item, which was difficult and non-discriminating, leaving two items in the excerpt.

TABLE 63 (contimued)

| Brcerpt | Source: Prelim. <br> Test, and. <br> Excerpt | Item | Level of Confldence in Discrimination | Level of Difificulty (\% Total group passing) |
| :---: | :---: | :---: | :---: | :---: |
| 11 | 1-5 | 1. | .01 | . 28 |
|  |  | 2 | . 01 | . 61 |
| 12 | 1-28 | 1 | - | . 56 |
|  |  | 2 | . 05 | .14 |
|  |  | 3 | (.10) | .58 |
|  |  | 4 | - | . 36 |
|  |  | 5 | . 01 | . 58 |
| 13 | 1-41 | 1 | . 01 | . 22 |
|  |  | 2 | .01 | . 39 |
|  |  | 3 | .01 | . 81 |
|  |  | 4 | , | . 47 |
| 14 | 1-42 | 1 | $\square$ | . 14 |
|  |  | 2 | . 05 | .11 |
|  |  | 3 | .01 | .17 |
|  |  | 4 | . 05 | . 114 |
|  |  | 5 | $\underline{-1}$ | .14 |
|  |  | 6 | . 01 | . 44 |
| 15 | 1-43 | 1 | . 01 | .39 |
|  |  | 2 | .05 | . 33 |
|  |  | 3 | . 01 | . 50 |
|  |  | 4 | .01 | .56 |
| 16 | 1-44 | 1 | . 01 | .47 |
|  |  | 2 | - | .47 |
|  |  | 3 | .01 | .53 |
|  |  | 4 | . 01 | .61 |
|  |  | 5 | . 05 | . 11 |
|  |  | 6 | $\square$ | .11 |
|  |  | 7 | . 05 | . 67 |
| 17 | 1-45 | 1 | . 01 | 147 |
|  |  | 2 | . 02 | . 25 |
|  |  | 3 | - | .67 |
|  |  | 4 | $\cdots$ | . 39 |
|  |  | 5 | .01 | . 44 |
|  |  | 6 | .05 | .19 |
|  |  | 7 | . 01 | .53 |
|  |  | 8 | . 01 | .56 |
|  |  | 9 | . 05 | . 11 |
|  |  | 10 | .01 | .19 |
|  |  | 11 | . 05 | .22 |
|  |  | 12 | . 01 | . 25 |
|  |  | 13 | . 01 | .53 |

## TABEE 63 (continued)

| Excerpt | Source: Prelim. <br> Test and <br> Excerpt | Item | Level of Confidence in Disorimu ination | Jevel of Difificulty (\% Total group passing) |
| :---: | :---: | :---: | :---: | :---: |
| 18 | 1-24 | 1 | . 01 | .47 |
|  |  | 2 | (.06) | .25 |
|  |  | 3 | - | . 50 |
| 19 | 2-11 | 1 | (.06) | . 61 |
|  |  | 2 | (.10) | . 46 |
|  |  | 3 | . 01 | .29 |
|  |  | 4 | - | . 39 |
| 20 | 1-37 | 1 | . 01 | . 42 |
|  |  | 2 | .01 | .39 |
|  |  | 3 | . 05 | -22 |
| 21 | 2-10 | 1 | . 01 | . 54 |
|  |  | 2 | . 05 | .43 |
|  |  | 3 | .01 | .39 |
|  |  | 4 | - | .25 |
|  |  | 5 | (.10) | .36 |
| 22 | $2-14$ | 1 | . 05 | . 25 |
|  |  | 2 | :01 | .64 |
|  |  | 3 | - | . 29 |
|  |  | 4 | (10) | .54 |
| 23 | 1-25 | 1 | . 01 | .56 |
|  |  | 2 | - | .17 |
|  |  | 3 | .05 | . 28 |
|  |  | 4 | . 01 | . 28 |
|  |  | 5 | .01 | . 33 |
|  |  | 6 | .05 | -28 |
| * +24 | 1-35 | 1 | .01 | .69 |
|  |  | 2 | (.06) | .25 |
|  |  | 3 | .01 | .25 |
| 25 | 1-17 | 1 | .05 | -31 |
|  |  | 2 | . 05 | -22 |
|  |  | 3 | - 05 | . 50 |
| 26 | 1-27 | $\frac{1}{2}$ | .05 .05 | . 37 |
|  |  | 3 | - | . 44 |

*erhis excerpt was revised and re-recorded to elininate the original fourth itam, which discriminated negatively.

TABLE 63 (concluaed)

| Exceryt | Source Prelin. Teest and Ezcerpt | Itom | Level of Confidence in Discrintination | Level of Duif ficulty (\% Total Group passing) |
| :---: | :---: | :---: | :---: | :---: |
|  | 1-3 | 1 | . 05 | 114 |
|  |  | 2 | . 05 | If |
|  |  | 3 | .05 | .19 |
| 28 | 1-4 | 1 | (.06) | .25 |
|  |  | 2 | . 01. | . 67 |
| 29 | 1240 | 1 | 01 | . 31 |
|  |  | 2 | . 01 | .28 |
|  |  | 3 | . 01 | .17 |
|  |  | 4 | . .01 | .17 |
|  |  | 5 | (.06) | .39 |
|  |  | 6 | . 05 | .33 |
|  |  | 7 | . 01 | .25 |
| 30 | 1-46 | 1 | . 01 | .25 |
|  |  | 2 | . .05 | 0.11 |

茾为This excerpt was revised to eliminate the ofiginal first item and add the same type of error as in ttem 2. This was possible since the notes and context were identical. The difficulty and discriminating pover of the item were then assumed to be similar to that of itam 2.

FIMAL TAST FORU ANALYSIS OF THE QUALITI AHD FTNCTION OF THE
 THTHRALS, ACCORDMG TO THEIR WOTATIOH, IMTO \%HICH PERFORUANCE ERHORS WIRE INTRODUCED

| Bxoexpt | Item | Intervals | Triads |  | 7th Chords |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  |  | Melodic Harmonic | Quazity | Function | Quality | Function |
| 1 | 1 | P4 |  |  |  |  |
|  | 2 |  | $\mathrm{m}_{1}$ | III |  |  |
|  | 3 |  | $\mathrm{m}_{1}$ | VI |  |  |
| 2 | 1 | P4 |  |  |  |  |
|  | 2 |  | m | VI |  |  |
|  | 3 |  | In | III |  |  |
| 3 | 1 |  | H | IV1 |  |  |
| 4 | 1 |  | mi | I |  |  |
|  | 2 |  | m | I |  |  |
|  | 3 |  | 313 | IVI |  |  |
|  | 4 |  | mi | I |  |  |
|  | 5 |  | m | I |  |  |
| 5 | 1 |  |  |  | D | V |
|  | 2 |  | \% | $I$ |  |  |
|  | 3 |  |  |  | $\mathrm{m}_{1}$ | II |
|  | 4 |  | \% | I |  |  |
|  | 5 |  | M | IV | III | II |
| 6 | 1 |  | M | IV |  |  |
| 7 | 1 |  | $m_{1}$ | III |  |  |
|  | 2 |  | 䐴 | $I_{2}$ |  |  |
|  | 3 | m 2 |  |  |  |  |
| 8 | 1 |  | $m_{1}$ | $\nabla_{1}$ |  |  |
|  | 2 |  | m | I |  |  |
|  | 3 |  | $\mathrm{M}_{1}$ | VII |  |  |
|  | 4 |  | m | さ |  |  |

Key to symbols: $\quad$ - major; $n$ - miner; $P$ - perfect; $d-$ diminished; $D$ - dominant; $X$ - chord th doninant function; subscripts 1,2 or 3 first, second, or third inversion.

Thais 64 (continued)


TABIE 64 (contimed)


## TABLE 64 (concluad )


*Possible analysis: in quality, diminushed 7th chord in 3rd inversion, with an appogiaturas in function, a dininished 7th chord of borromed dominant function, 3rd inversion, with appogiatura.

TABLS 65

FIMAL TEST: SCORES OBLALNED TOR SPLT HALUES OF TEST,
 TESS RELTABTLITY

| $\begin{gathered} \text { odd } \\ \mathbf{X} \end{gathered}$ | Even I | $\frac{\text { Odd }}{X}$ | Bren I |
| :---: | :---: | :---: | :---: |
| 10 | 13 | 8 | 9 |
| 10 | 14 | 29 | 28 |
| 7 | 9 | 24 | 27 |
| 18. | 16 | 43 | 50 |
| 29 | 30 | 18 | 13 |
| 12 | 13 | 18 | 20 |
| 13 | 17 | 17 | 23 |
| 14 | 12 | 43 | 50 |
| 8 | 12 | 38 | 40 |
| 16 | 17 | 24 | 21 |
| 8 | 8 | 4 | 4 |
| 13 | 17 | 18 | 19 |
| 36 | 33 | 5 | 11. |
| 14 | 17 | 10 | 114 |
| 12 | 14 | 33 | 33 |
| 36 | 38 | 46 | 43 |
| 18 | 17 | 49 | 49 |
| 12 | 13 | 42 | 42 |
| 47 | 51 | 33 | 31 |
| 10 | 18 | 32 | 26 |
| 34 | 35 | EXP1090 | $\Sigma Y=1151$ |
| 22 | 27 |  |  |
| 11 | 15 |  |  |
| 27 | 25 |  |  |
| 12 | 11. |  |  |
| 29 | 24 |  |  |
| 11 | 17 |  |  |
| 31 | 31 |  |  |
| 13 | 20 |  |  |
| 15 | 15 |  |  |

## APPENDTX $F$

## Letter to Heads of Husic Depariments

## 

One of the inportant activities of masicians, particulariy of those who direot musical ensembles, is to make sure that the perfomance of music agrees with the printed score in ths melodic and harmonic aspeots. In order to detemine the levels of the particular ability Which existe mong musicians, and to study the relationship of certain factors to thid ablitity, I have devised a test and brief accompanying questionnaire. Your Interest and help in thoir administration to your musio students mould be greatly appreciated.

The test consisto of a number of excerpts in score of representam tive choral muste which are to be inspected as they are heard in perfomance by means of a magnetic tape recording. In the performance certain errors are introcuced. such as might occur due to inaccurate sight reading by the singers. Any discrepancies between score and performance are to be indicated by a sinple system of marking on the score. Adminn istration of the test will require no more than one class period approximately ten minutes for filling out the questionnaire and reading the test instmuctions, and thirty-six minutes for the test itself.

From this brief description it is hoped that you will be intorm osted in administering the test to your musio students. Infomation is desired regarding porionance on this test by all levels of your music students from freshnen to seniors, and also graduate students. Since the test may measure the extent of a functional use of knowledge of melodic intervals and hamonic structure (the test includes a representative sampling of frequently used chords), it will be greatiy
apprectated if the test could be administered to first and second year muste theory classes, to upper level theory classes, and to conducting classes.

Please indicate on the enclosed form the requested information. For your convenience a self-addressed stamped envelope is anclosed. The number of tests you request, together with the reel of tape, will be mailed to you in the near future. All individual test data will of course be confidential. A sumary of results when completed will be provided at your request.

Please note that in order to adminster the test you should have available for use a magnetic tape playing machine with a speed of $7 \frac{7}{3}$ inohes per second.

Anticipating and thanking you for your kind cooperation, I an Very sincerely yours,
(Signature)
Encl.


[^0]:    Tincent jones. 解ic Equcation in the college. C. C. Birchard and Co. Foston, 2049. p. 59.

    2The tem "abilsty" is used in a general sense as "that which one has attaired through troinins, practice or experience." (thax Schoeng The Psycholosy of husic. The Ronala Press Company, Her Tonts 1940. pp. I6I162.)

[^1]:    $I_{\text {Jones, }}$ p. 57.
    2uarie Agnow. "A Comparison of the Auditory Images of Kusicians, Psycholocists, and Children." Psychological ionormaphs, 31, 1 (1922) Thole ino. $140,268-273$.

[^2]:    Itarie Agnem. "The Audatory Imasery of Great Composers." Psycholocical Honograbhs, 31, 1 (1922) Whole No. 110, 279-237.
    anoger Sessions. The Husical Experionce of Oomposor, Fexformor, Estoner. Pxincoton Unversity Press, Princetong UoJ. 1950. P. 29.
    $3_{\text {Robert }}$ A. Belcher. "The status and Function of Ear Training in herican Colleees and Coneervatories." Mra Proccedinss, 33 (1938) 192.

[^3]:    ${ }^{\text {ISamuel T. Burns. Hamonic Skills Used by Selected Hich School }}$ Choral Leaders. Dureau of Publicathons, Teachers College. Colunbia Thiversity, New xoric, 1945. P. $95^{\circ}$.

[^4]:    Tharara Bo, Greene. Measurenents of Human Rehavior. Tho Odyssy Press, Nen Tork, 1952. P. 306.

[^5]:    11．W．Torgerson and 定mest Fahnestack．Torgerson－Fahnestook Muaic Tests．Publia School Publishing Co．，Bloomington，I11．，1926．

    2Herbert Hutchinson．Hutehinson Music Tests．Publio School Publishing Coo，Bloomington，Ill．，1924．

    Brank A．Deach and H．Do Schrammel．Beach Ausic Testes，Kamual of Dtwections．Dureau of Educational teasurements，Kansas State Teachers College，Eriporia，Kansas，1920，rev． 1930.

[^6]:    Irichard D. Allen, falter fi. Batterfield, and Marguertte Tully.
     Company, Yonkers, lfow Xoric, 1932.

    2Glenn Gildersleeve and Woyn Soper. Mosical Achievernent Test. Burgau of Publications, Teachers College, Columbit Tniversity, New Yorle 1933.

    3waldo C. Panloratz. "A Comparison of Sono Group Eustcal Achievement Tests as Related to I.Q. and Teacher Rating." Unpublished master's thosis, thivercity of Romsan, 1950. P. 21.

[^7]:     Company, Yonkers, Hem Iork 1923.

[^8]:     Pgychological Monofraphs, 39, 2 (1028) tholc No. 178 , 21.5-156.

    2T. S. Solisbury and i. N. Soith. "Proenosis of Sight-sineine Abllity of Momal school stacionts." Journol of Appliod Psycholocy, 13 (1929) $425-439$.
    $3_{\text {Adolph Ottorstein and haynond M. Wosher. } 0-4 \text { Sight-Singing Test. }}$ Stanford univorsity Press, palo Alto, Califormiag 1932.

    4pamond H. Moshor. A Guady of the Group Rethod of Leesurement of Sifht-stncing. Burear of publications, feachers College, Columia thiversity, Jew Tork, 1925.

[^9]:    Hiosher, p. 5h. This mould then be true also for hamonic dictethong but would be more difficult to test.

    Frances inimit. "The Correlation between Achievement anci Canacity in Music." Jommal of Edncational Research, 17 (1923) 50 56.

[^10]:    IJanes Aliferis and J. E. Stecklein. "The Devoloment oit a College Entrance Test in Nusic Achievenent." Joumal of Rosearch in Jusic Educabion, 1, No. 2 (Fall 1953) 83m6. The tost is available from Univeroity of Mimesota Proso, 保meapolis, Mim.

[^11]:    Thax Schoen. The Psycholory of Yusic. The Ronald Press Company, Wew Tork. 1940. P. 161.

[^12]:     Ecucators Journal, 24.3 Ko. 3 (1937) 25-26.

[^13]:    $1_{\text {R. Famsworth. "Studies in the Psychology of Tone." Genetic }}$ Psychology Monographs, 15:1 (1934) 1-91.

    2James I. 倓rell. The Psychology of Music. W. W. Morton and Company, Inc., New York, 1937. P. 323f.

[^14]:    R.M. Draice. "The Validity and Reliability of Tests of Masioal Talent." Journal of Applied Psychology, 17 (1933) 447-458.
    ${ }^{2}$ Drake Kusical Memory Testy A Tost of Musical Talent. Public School Publishing Co., Bloomington, I11., 1934.

    3T. H. Madison. "Interval Discrimination as a Measure of Musical Aptitude." Archives of Psychology, 1942, Ho. 206, 1-99. The test is obtainable from the author, University of Indiana; Bloonington, Indiana.

[^15]:    $I_{\text {Max }}$ Schoen. "Tests of Rusical Feeling and \%usical Understandinge" Journal of Gommerative Psycholocys, 5 (1925) 31-52.

[^16]:    ${ }^{1}$ Drake. "The Validity and Reliability of Tests of wusical Talent."
    $2_{\text {Robert }}$ w. Lindin. "The Development and Validation of a Set of孜usical Ability Tests." Psychological Monographs, 63, 10 (1949) Whole No. 305, 1-20.

[^17]:    $I_{\text {Reported by Schoen in The Psychology of Music, Pp, 180-181 and }}$ 189-190.
    2. A. Petran. "An Analysis of the Ortmann Tests in Tonel Discrimination and Hemory." Bulletin of the Anerican fastcological Societys 25, 2 ( 1937 ) 304-306.

[^18]:    $I_{\text {E. Thayer Gaston. A Test of }}$ Husicality. Third Edition, Revised.
    

[^19]:    1. Lowery " Cadance and Ehrase Testa in Music." British Journel of Psychology, 17 (1926) 111-118.

    2J. Kraimasser. "Tests ana Heasuremsnts in Music." Psychological Fulletin, 25 (1928) 284-301. Hwalwasser Test of Helodic and Hamonic Sensitivity. Victor Talleing Machine Co., Canden, Na J. 1926.
    ${ }^{3}$ Kate Hevner and John Lansbury. Oregon Yasic Discrintnation Test. O. H. Stoelting Co., Chicago, 1935.

[^20]:    $1_{\mathrm{K}}$. Hevner. "Tests for Aesthetic Appreciation in the Field of Fusic." Joumel of Applisa Parohology, 14 (3930) 470-77.
    $2_{\text {ra }}$. Hevner. "A study of Teste for the Appreciation of Husic." Journal of Applied Psychology, 15 (1931) 575-83.

[^21]:    10. C. Hoss. Kaasurement in Today is schoois. Frentice-Hall, Inc., Nerr Yorik, 1948. P. 76.
[^22]:    $I_{\text {Samuel T. Burns. Harmonic Skills Used by Selected Migh School }}$ Choral Leaders. Bureau of Publications, Teachers College, Columbia University, New York, 1945.
    ${ }^{2}$ Irara G. Hoggard. Improving Music Reading in the Gharal Rehearsal. The Author, New York, 1947.

[^23]:    $I_{\text {Adapted }}$ srom Hoggard, Fig. $2 ;$ p. 3.

[^24]:    IA number of such errors, as a matter of fact, actually occurred in the author's own experience.

[^25]:    $1_{\text {Arthur E. Traxler. "Administering and Scoring the Objective }}$ Test." Eduoational Measurement. E.F. Iindquist, editor. American Council on Education, Washington, D. C., 1951. P. 365.

[^26]:    ${ }^{1}$ Dorothy C. Adkins and others. Construction and Analysis of Achievement Tests. U.S. Government Printing Office, Washington, 1947. P. 151.
    ${ }^{2}$ J.P. Guilford. Fundamental Statistics in Psychology and Education. Hedraw-Hill Book Co., New York, 1950. P. 157.

    3 Ibid. pp. 492-493.
    4 Ibid., pp. 159.

[^27]:    Irederick B. Davis. "Iten Selection Techniques." Educational Measurement. E. F. Iindquist, editor. American Council on Edacation, Wasingtong. D. ©., 1951. Of. pp. 267-285.

[^28]:    $Z_{\text {Guilford, }}$ p. 499.

[^29]:    Key to symbols: ${ }^{(4-m a j o r, ~ m-m i n o r ; ~ d-d i m i n i s h e d . ~ P-p e r f e c t, ~ D-i o m i n a n t . ~}$ *Quality: 3rd inversion of dininishod 7th chord, with appogiaturay Funcm tion, incomplete 13th chord, 3rd inversion, of dominant function.

[^30]:    1J. P. Guilford. Fundamental Statistics in Psychology and Education. MeGrawmitil Book Co., Hew York, 1950. P. 279.

[^31]:    IGuilford, p. 228.
    2Palner 0. Johnsone Statistical Methods in Research. PrenticeHall, Inc., New York, 1949. Pp. 73-75.

    $$
    3_{\text {Guilford, pp. }} \text { 232-233. }
    $$

[^32]:    $1_{\text {Renneth }}$ E, Anderson, $\mathrm{E}_{\text {, Thayer Gaston, and Melvin H. Zack. }}$ A Kothod of Establishing Test Validity. Unpublished paper, University of Kansas, Lawrence, 1954, To be published in the Karch, 1955 issue of Journal of Experinental Education.

[^33]:    —

    * "The Velidfty and Reliability of Tests of Misical Talent." Journal of Applied Psycholozy, 17, 447-458. 1933.

