AN EXPERIMENTAL INVESTIGATION OF
THE EFFECTIVENESS OF TRAINING
ON ABSOLUTE PITCH IN ADULT MUSICIANS

DISSERTATION

Presented in Partial Fulfillment of the Requirements for
the Degree Doctor of Philosophy in the Graduate
School of the Ohio State University

By

Mark Alan Rush, B. A., B. M., M. M.

* * * * *

The Ohio State University
1989

Dissertation Committee:

David Butler
A. Peter Costanza
Burdette L. Green
Mari Riess Jones

Approved by

Advisor
School of Music
Copyright by
Mark A. Rush
1989
"Where there is no experiment there can be no science."

Carl E. Seashore,
in
The Inheritance of Musical Talent
(1920)
To the memory of my Grandfather,

Neil D. Rush
1903—1988

Who provided me with innumerable opportunities
and showed me the value of hard work.
ACKNOWLEDGMENTS

I wish to express my gratitude to the members of the dissertation committee. The advice and counsel of Professors Peter Costanza and Mari Riess Jones was appreciated, particularly in the area of data analysis. I am especially indebted to Professor Burdette Green, whose acute eye and sensitive pen improved this document in many ways. Lastly, I would like to express my gratitude to my advisor, Professor David Butler, not only for his help in the preparation of this study, but for his direction and friendship throughout my graduate education.

Special thanks go to Professor William Loadman, who assisted in the experimental design, and to Professor Gregory Proctor, whose expertise as a translator proved invaluable. The help of the Ohio State University Libraries staff was extremely important: particularly that of Scott Seaman and Debbie Cameron, of the Interlibrary Loan Department; Dr. Thomas Heck, of the Music and Dance Library; and Marty Jamison, of the Education and Psychology Library.

Without the receipt of a Graduate Student Alumni Research Award and the support of a Presidential Fellowship, this project could not have been undertaken. Special thanks go to the subjects who voluntarily gave their time to this study.

Finally, to my wife Dee, without whose patience and encouragement this research would never have been completed.
VITA

September 8, 1958 ........................................ Born - Columbus, Ohio

1981 .......................................................... Bachelor of Arts, Music
The Ohio State University
Columbus, Ohio

1982 .......................................................... Bachelor of Music, Jazz Studies
The Ohio State University
Columbus, Ohio

1983 .......................................................... Master of Music, Music Composition
The Ohio State University
Columbus, Ohio

1983 - 1987 .................................................. Graduate Teaching Associate
Theory and Composition Division
The Ohio State University
School of Music

1984 - 1987 .................................................. Graduate Research Associate
Music Perception Studio
The Ohio State University
School of Music

1987- 1988 .................................................. Presidential Fellow
The Ohio State University
PUBLICATIONS


PRESENTATIONS


FIELDS OF STUDY

Major Field: Music Theory

Studies in Music Perception and Cognition: Dr. David Butler
Studies in Music Analysis: Dr. Gregory Proctor, Dr. William Poland
Studies in History of Music Theory: Dr. Burdette Green
Studies in Electronic Music: Dr. Thomas Wells
# TABLE OF CONTENTS

AKNOWLEDGMENTS ........................................................................ iii

VITA ......................................................................................... iv

LIST OF TABLES .......................................................................... ix

LIST OF FIGURES ......................................................................... xii

<table>
<thead>
<tr>
<th>CHAPTER</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>I. INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>Background</td>
<td>1</td>
</tr>
<tr>
<td>Definitions of Terms</td>
<td>2</td>
</tr>
<tr>
<td>Description of the Problem</td>
<td>2</td>
</tr>
<tr>
<td>Statement of Purpose</td>
<td>4</td>
</tr>
<tr>
<td>Need for the Study</td>
<td>4</td>
</tr>
<tr>
<td>Limitations of the Study</td>
<td>5</td>
</tr>
<tr>
<td>II. REVIEW OF LITERATURE</td>
<td>6</td>
</tr>
<tr>
<td>Introduction</td>
<td>6</td>
</tr>
<tr>
<td>Meyer (1899)</td>
<td>7</td>
</tr>
<tr>
<td>Jadassohn (1899)</td>
<td>10</td>
</tr>
<tr>
<td>Köhler (1915)</td>
<td>11</td>
</tr>
<tr>
<td>Copp (1916)</td>
<td>13</td>
</tr>
<tr>
<td>Anonymous (1916)</td>
<td>15</td>
</tr>
<tr>
<td>Maryon (1919)</td>
<td>16</td>
</tr>
<tr>
<td>Gough (1922)</td>
<td>19</td>
</tr>
<tr>
<td>Mull (1925)</td>
<td>22</td>
</tr>
<tr>
<td>Mal'tseva (1925)</td>
<td>29</td>
</tr>
<tr>
<td>Wedell (1934)</td>
<td>31</td>
</tr>
<tr>
<td>Komatsu (1940)</td>
<td>37</td>
</tr>
<tr>
<td>Hindemith (1949)</td>
<td>40</td>
</tr>
<tr>
<td>Pollack (1952)</td>
<td>42</td>
</tr>
<tr>
<td>Reference</td>
<td>Page</td>
</tr>
<tr>
<td>-------------------------</td>
<td>------</td>
</tr>
<tr>
<td>Hartman (1954)</td>
<td>46</td>
</tr>
<tr>
<td>Deutsch (1954)</td>
<td>49</td>
</tr>
<tr>
<td>Lundin &amp; Allen (1962)</td>
<td>52</td>
</tr>
<tr>
<td>Lundin (1963)</td>
<td>57</td>
</tr>
<tr>
<td>Terman (1965)</td>
<td>58</td>
</tr>
<tr>
<td>Fisher (1965)</td>
<td>61</td>
</tr>
<tr>
<td>Vianello (1966)</td>
<td>63</td>
</tr>
<tr>
<td>Cuddy (1968)</td>
<td>65</td>
</tr>
<tr>
<td>Brady (1970)</td>
<td>74</td>
</tr>
<tr>
<td>Cuddy (1970)</td>
<td>79</td>
</tr>
<tr>
<td>Cuddy (1971)</td>
<td>85</td>
</tr>
<tr>
<td>Heller &amp; Auerbach (1972)</td>
<td>91</td>
</tr>
<tr>
<td>Falk (1975)</td>
<td>94</td>
</tr>
<tr>
<td>Hurni-Schlegel &amp; Lang (1978)</td>
<td>96</td>
</tr>
<tr>
<td>Waters (1980)</td>
<td>99</td>
</tr>
<tr>
<td>Grebel'nik (1984)</td>
<td>106</td>
</tr>
<tr>
<td>Faiivre (1986)</td>
<td>111</td>
</tr>
<tr>
<td>Summary</td>
<td>117</td>
</tr>
</tbody>
</table>

III. PROBLEM

- The Role of Chroma in Absolute Pitch
- Selection of the Training Method
- Burge's Technique
- Hypotheses

IV. METHODOLOGY

- Experimental Design
- Subjects
- Description of the Pretest and Posttest
- Description of the Training Procedure

V. RESULTS AND DISCUSSION

- Interpretation of Test Results
- Discussion of Results Obtained from the Complete Experimental and Control Groups
- Divisions of the Experimental Group
- Discussion of Results Obtained from Group C
- Discussion of Results Obtained from Group B
- Discussion of Results Obtained from Group A
- The Correlational Method
<table>
<thead>
<tr>
<th>VI. SUMMARY</th>
<th>207</th>
</tr>
</thead>
<tbody>
<tr>
<td>Internal and External Validity</td>
<td>207</td>
</tr>
<tr>
<td>Generalizability of the Results of the Study</td>
<td>208</td>
</tr>
<tr>
<td>Conclusions</td>
<td>209</td>
</tr>
<tr>
<td>Implications</td>
<td>213</td>
</tr>
<tr>
<td>Recommendations for Future Study</td>
<td>214</td>
</tr>
</tbody>
</table>

| BIBLIOGRAPHY | 216 |

<table>
<thead>
<tr>
<th>APPENDICES</th>
<th>252</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. Materials Relative to Chapter IV</td>
<td>252</td>
</tr>
<tr>
<td>Questionnaire</td>
<td>253</td>
</tr>
<tr>
<td>Test Music</td>
<td>254</td>
</tr>
<tr>
<td>Test Directions</td>
<td>259</td>
</tr>
<tr>
<td>Consent Form</td>
<td>260</td>
</tr>
<tr>
<td>Sample Practice Log</td>
<td>261</td>
</tr>
<tr>
<td>B. Individual Error Analyses by Pitch Class</td>
<td>262</td>
</tr>
<tr>
<td>C. Individual Error Analyses by Interval Class</td>
<td>315</td>
</tr>
<tr>
<td>D. Training Log Summaries</td>
<td>368</td>
</tr>
<tr>
<td>E. Experimental Group Subjects' Comments</td>
<td>396</td>
</tr>
</tbody>
</table>
### LIST OF TABLES

<table>
<thead>
<tr>
<th>TABLE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Frequencies of tuning fork tones and the octave relationships between them in the Meyer and Heyfelder experiment</td>
<td>8</td>
</tr>
<tr>
<td>2. Frequencies of tuning fork tones in the Meyer and Heyfelder experiment reduced to one octave and shown in ascending order. Boldface numbers indicate the frequencies of the tones actually present</td>
<td>8</td>
</tr>
<tr>
<td>3. All colors (shown with color names and frequencies in Angstroms) and tones (shown with pitch labels and frequencies in Hertz) in the order in which the Marcotone system states they are to be learned</td>
<td>18</td>
</tr>
<tr>
<td>4. Summary of individual subjects’ data for Lundin &amp; Allen (1962) showing a) the percentage of correct responses on the 24-tone pretest, b) the percentage of correct responses on the 24-tones posttest, c) the percentage correct responses during the first four training sessions, d) the percentage correct responses during the last four training sessions, e) the percentage improvement from the first four training sessions to the last four training sessions, f) the average error in semitones for the first four sessions, g) the average error in semitones for the last four sessions, and h) the total number of training sessions each subject attended</td>
<td>56</td>
</tr>
<tr>
<td>5. The response labels associated with frequencies, in Hertz, of stimulus tones utilized by Vianello (1966)</td>
<td>64</td>
</tr>
<tr>
<td>6. The four possible training scenarios in Cuddy’s Experiment 2 (1968)</td>
<td>69</td>
</tr>
<tr>
<td>7. The four possible testing scenarios in Cuddy’s Experiment 2 (1968)</td>
<td>70</td>
</tr>
</tbody>
</table>
8. The four possible training scenarios in Cuddy's Experiment 3 (1968) ................................................................. 73

9. The percentage of correct judgments and bits of information transmitted, averaged across both Set A and Set B, for Cuddy's (1970) Experiment ................................................................. 84


11. The order and arrangement of the 12 stimuli presented at each of the nine training phases in Faivre (1986) .................................................. 114

12. Individual background data for the members of the experimental group ................................................................. 138

13. Individual background data for the members of the control group ................................................................. 140

14. Mean number of correct responses (120 possible), standard deviation, and number of observations on the pretest and the posttest for the experimental and control groups ...................... 154

15. Mean absolute error (semitones), standard deviation, and number of observations on the pretest and the posttest for the experimental and control groups .................................................. 156

16. Mean amount of information transmitted (bits), standard deviation, and number of observations on the pretest and the posttest for the experimental and control groups ...................... 158

17. Mean number of correct responses (120 possible), mean absolute error (semitones), mean amount of information transmitted (bits), their standard deviations, and number of observations on the pretest and the posttest for Group C and the control group .... 173

18. Mean number of correct responses (120 possible), mean absolute error (semitones), mean amount of information transmitted (bits), their standard deviations, and number of observations on the pretest and the posttest for Group B and the control group .... 183

19. Mean number of correct responses (120 possible), mean absolute error (semitones), mean amount of information transmitted (bits), their standard deviations, and number of observations on the pretest and the posttest for Group A and the control group .... 194
20. Summary of the total number of training sessions, their combined duration, the highest difficulty level practiced during the final week, and the week of the experiment in which it occurred for the subjects in the experimental group .................................. 395
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>FIGURE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. An example of the order of memorization for a series of tones, arbitrarily beginning on Eb, recommended by Deutsch (1954)</td>
<td>51</td>
</tr>
<tr>
<td>2. Deutsch's three-tone groups to be memorized in a simulated form of absolute pitch by (a) sopranos, (b) altos, (c) tenors, and (d) basses</td>
<td>52</td>
</tr>
<tr>
<td>3. The musical notation, note names, and frequencies in Hertz for the 12 pitches used in Cuddy's Experiment 1 (1968)</td>
<td>67</td>
</tr>
<tr>
<td>4. The relative spacing of Set A and Set B in Cuddy's (1970) Experiment</td>
<td>81</td>
</tr>
<tr>
<td>5. Digitally enhanced reproduction of Drobisch's helical representation of musical pitch (1855)</td>
<td>126</td>
</tr>
<tr>
<td>6. The effect of training on the mean number of correct responses for the experimental and control groups</td>
<td>155</td>
</tr>
<tr>
<td>7. The effect of training on the mean absolute error for the experimental and control groups</td>
<td>157</td>
</tr>
<tr>
<td>8. The effect of training on the mean amount of information transmitted for the experimental and control groups</td>
<td>159</td>
</tr>
<tr>
<td>9. Mean number of correct responses, mean absolute error, and mean amount of information transmitted as a function of pitch class for the subjects in the experimental group</td>
<td>161</td>
</tr>
<tr>
<td>10. Mean number of correct responses, mean absolute error, and mean amount of information transmitted as a function of pitch class for the subjects in the control group</td>
<td>162</td>
</tr>
</tbody>
</table>
11. Mean number of responses as a function of error interval for the subjects in the experimental group ............................................ 164

12. Mean number of responses as a function of error interval for the subjects in the control group .................................................. 164

13. Mean number of responses as a function of absolute error interval for the subjects in the experimental group .................. 165

14. Mean number of responses as a function of absolute error interval for the subjects in the control group .......................... 165

15. Individual scores of the number of correct responses for the subjects in the experimental group ........................................ 167

16. Individual scores of the number of correct responses for the subjects in the control group .................................................. 167

17. Individual scores of the mean absolute error for the subjects in the experimental group .................................................. 168

18. Individual scores of the mean absolute error for the subjects in the control group .................................................. 168

19. Individual scores of the amount of information transmitted for the subjects in the experimental group .................. 169

20. Individual scores of the amount of information transmitted for the subjects in the control group ........................................ 169

21. Mean number correct, mean absolute error, and mean amount of information transmitted for subjects in Group C and the control group ........................................................................................................... 174

22. Mean number of correct responses, mean absolute error, and mean amount of information transmitted as a function of pitch class for the subjects in Group C .................................................. 176

23. Mean number of responses as a function of error interval for the subjects in Group C .................................................. 178

24. Mean number of responses as a function of absolute error interval for the subjects in Group C ........................................ 178
25. The effect of training on the individual scores of the number of correct responses, the mean absolute error, and the amount of information transmitted for the subjects in Group C ........................................ 180

26. Mean number correct, mean absolute error, and mean amount of information transmitted for subjects in Group B and the control group ................................................................................................................................. 184

27. Mean number of correct responses, mean absolute error, and mean amount of information transmitted as a function of pitch class for the subjects in Group B ........................................................................................................ 186

28. Mean number of responses as a function of error interval for the subjects in Group B ........................................................................................................................................... 188

29. Mean number of responses as a function of absolute error interval for the subjects in Group B ........................................................................................................................................... 188

30. The effect of training on the individual scores of the number of correct responses, the mean absolute error, and the amount of information transmitted for the subjects in Group B ........................................................................................................ 190

31. Mean number correct, mean absolute error, and mean amount of information transmitted for subjects in Group A and the control group ........................................................................................................ 195

32. Mean number of correct responses, mean absolute error, and mean amount of information transmitted as a function of pitch class for the subjects in Group A ........................................................................................................ 197

33. Mean number of responses as a function of error interval for the subjects in Group A ........................................................................................................................................... 199

34. Mean number of responses as a function of absolute error interval for the subjects in Group A ........................................................................................................................................... 199

35. The effect of training on the individual scores of the number of correct responses, the mean absolute error, and the amount of information transmitted for the subjects in Group A ........................................................................................................ 201

36. Number correct, mean absolute error, and information transmitted as a function of pitch class for Subject A in the experimental group ........................................................................................................................................... 263
37. Number correct, mean absolute error, and information transmitted as a function of pitch class for Subject B in the experimental group .......................................................... 264

38. Number correct, mean absolute error, and information transmitted as a function of pitch class for Subject C in the experimental group .......................................................... 265

39. Number correct, mean absolute error, and information transmitted as a function of pitch class for Subject D in the experimental group .......................................................... 266

40. Number correct, mean absolute error, and information transmitted as a function of pitch class for Subject E in the experimental group .......................................................... 267

41. Number correct, mean absolute error, and information transmitted as a function of pitch class for Subject F in the experimental group .......................................................... 268

42. Number correct, mean absolute error, and information transmitted as a function of pitch class for Subject G in the experimental group .......................................................... 269

43. Number correct, mean absolute error, and information transmitted as a function of pitch class for Subject H in the experimental group .......................................................... 270

44. Number correct, mean absolute error, and information transmitted as a function of pitch class for Subject I in the experimental group .......................................................... 271

45. Number correct, mean absolute error, and information transmitted as a function of pitch class for Subject J in the experimental group .......................................................... 272

46. Number correct, mean absolute error, and information transmitted as a function of pitch class for Subject K in the experimental group .......................................................... 273

47. Number correct, mean absolute error, and information transmitted as a function of pitch class for Subject L in the experimental group .......................................................... 274
48. Number correct, mean absolute error, and information transmitted as a function of pitch class for Subject M in the experimental group ................................................................. 275

49. Number correct, mean absolute error, and information transmitted as a function of pitch class for Subject N in the experimental group ................................................................. 276

50. Number correct, mean absolute error, and information transmitted as a function of pitch class for Subject O in the experimental group ................................................................. 277

51. Number correct, mean absolute error, and information transmitted as a function of pitch class for Subject P in the experimental group ................................................................. 278

52. Number correct, mean absolute error, and information transmitted as a function of pitch class for Subject Q in the experimental group ................................................................. 279

53. Number correct, mean absolute error, and information transmitted as a function of pitch class for Subject R in the experimental group ................................................................. 280

54. Number correct, mean absolute error, and information transmitted as a function of pitch class for Subject S in the experimental group ................................................................. 281

55. Number correct, mean absolute error, and information transmitted as a function of pitch class for Subject T in the experimental group ................................................................. 282

56. Number correct, mean absolute error, and information transmitted as a function of pitch class for Subject U in the experimental group ................................................................. 283

57. Number correct, mean absolute error, and information transmitted as a function of pitch class for Subject V in the experimental group ................................................................. 284

58. Number correct, mean absolute error, and information transmitted as a function of pitch class for Subject W in the experimental group ................................................................. 285
59. Number correct, mean absolute error, and information transmitted as a function of pitch class for Subject X in the experimental group ................................................................. 286

60. Number correct, mean absolute error, and information transmitted as a function of pitch class for Subject Y in the experimental group ................................................................. 287

61. Number correct, mean absolute error, and information transmitted as a function of pitch class for Subject Z in the experimental group ................................................................. 288

62. Number correct, mean absolute error, and information transmitted as a function of pitch class for Subject AA in the control group ................................................................. 289

63. Number correct, mean absolute error, and information transmitted as a function of pitch class for Subject BB in the control group ................................................................. 290

64. Number correct, mean absolute error, and information transmitted as a function of pitch class for Subject CC in the control group ................................................................. 291

65. Number correct, mean absolute error, and information transmitted as a function of pitch class for Subject DD in the control group ................................................................. 292

66. Number correct, mean absolute error, and information transmitted as a function of pitch class for Subject EE in the control group ................................................................. 293

67. Number correct, mean absolute error, and information transmitted as a function of pitch class for Subject FF in the control group ................................................................. 294

68. Number correct, mean absolute error, and information transmitted as a function of pitch class for Subject GG in the control group ................................................................. 295

69. Number correct, mean absolute error, and information transmitted as a function of pitch class for Subject HH in the control group ................................................................. 296
<table>
<thead>
<tr>
<th>Number</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>70.</td>
<td>Number correct, mean absolute error, and information transmitted as a function of pitch class for Subject II in the control group</td>
<td>297</td>
</tr>
<tr>
<td>71.</td>
<td>Number correct, mean absolute error, and information transmitted as a function of pitch class for Subject JJ in the control group</td>
<td>298</td>
</tr>
<tr>
<td>72.</td>
<td>Number correct, mean absolute error, and information transmitted as a function of pitch class for Subject KK in the control group</td>
<td>299</td>
</tr>
<tr>
<td>73.</td>
<td>Number correct, mean absolute error, and information transmitted as a function of pitch class for Subject LL in the control group</td>
<td>300</td>
</tr>
<tr>
<td>74.</td>
<td>Number correct, mean absolute error, and information transmitted as a function of pitch class for Subject MM in the control group</td>
<td>301</td>
</tr>
<tr>
<td>75.</td>
<td>Number correct, mean absolute error, and information transmitted as a function of pitch class for Subject NN in the control group</td>
<td>302</td>
</tr>
<tr>
<td>76.</td>
<td>Number correct, mean absolute error, and information transmitted as a function of pitch class for Subject OO in the control group</td>
<td>303</td>
</tr>
<tr>
<td>77.</td>
<td>Number correct, mean absolute error, and information transmitted as a function of pitch class for Subject PP in the control group</td>
<td>304</td>
</tr>
<tr>
<td>78.</td>
<td>Number correct, mean absolute error, and information transmitted as a function of pitch class for Subject QQ in the control group</td>
<td>305</td>
</tr>
<tr>
<td>79.</td>
<td>Number correct, mean absolute error, and information transmitted as a function of pitch class for Subject RR in the control group</td>
<td>306</td>
</tr>
<tr>
<td>80.</td>
<td>Number correct, mean absolute error, and information transmitted as a function of pitch class for Subject SS in the control group</td>
<td>307</td>
</tr>
<tr>
<td>Number</td>
<td>Title</td>
<td>Page</td>
</tr>
<tr>
<td>--------</td>
<td>----------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>81.</td>
<td>Number correct, mean absolute error, and information transmitted as</td>
<td>308</td>
</tr>
<tr>
<td></td>
<td>a function of pitch class for Subject TT in the control group</td>
<td></td>
</tr>
<tr>
<td>82.</td>
<td>Number correct, mean absolute error, and information transmitted</td>
<td>309</td>
</tr>
<tr>
<td></td>
<td>as a function of pitch class for Subject UU in the control group</td>
<td></td>
</tr>
<tr>
<td>83.</td>
<td>Number correct, mean absolute error, and information transmitted</td>
<td>310</td>
</tr>
<tr>
<td></td>
<td>as a function of pitch class for Subject VV in the control group</td>
<td></td>
</tr>
<tr>
<td>84.</td>
<td>Number correct, mean absolute error, and information transmitted</td>
<td>311</td>
</tr>
<tr>
<td></td>
<td>as a function of pitch class for Subject WW in the control group</td>
<td></td>
</tr>
<tr>
<td>85.</td>
<td>Number correct, mean absolute error, and information transmitted</td>
<td>312</td>
</tr>
<tr>
<td></td>
<td>as a function of pitch class for Subject XX in the control group</td>
<td></td>
</tr>
<tr>
<td>86.</td>
<td>Number correct, mean absolute error, and information transmitted</td>
<td>313</td>
</tr>
<tr>
<td></td>
<td>as a function of pitch class for Subject YY in the control group</td>
<td></td>
</tr>
<tr>
<td>87.</td>
<td>Number correct, mean absolute error, and information transmitted</td>
<td>314</td>
</tr>
<tr>
<td></td>
<td>as a function of pitch class for Subject ZZ in the control group</td>
<td></td>
</tr>
<tr>
<td>88.</td>
<td>Number of responses as a function of error interval for Subject A</td>
<td>316</td>
</tr>
<tr>
<td></td>
<td>in the experimental group</td>
<td></td>
</tr>
<tr>
<td>89.</td>
<td>Number of responses as a function of absolute error interval for</td>
<td>316</td>
</tr>
<tr>
<td></td>
<td>Subject A in the experimental group</td>
<td></td>
</tr>
<tr>
<td>90.</td>
<td>Number of responses as a function of error interval for Subject B</td>
<td>317</td>
</tr>
<tr>
<td></td>
<td>in the experimental group</td>
<td></td>
</tr>
<tr>
<td>91.</td>
<td>Number of responses as a function of absolute error interval for</td>
<td>317</td>
</tr>
<tr>
<td></td>
<td>Subject B in the experimental group</td>
<td></td>
</tr>
<tr>
<td>92.</td>
<td>Number of responses as a function of error interval for Subject C</td>
<td>318</td>
</tr>
<tr>
<td></td>
<td>in the experimental group</td>
<td></td>
</tr>
<tr>
<td>93.</td>
<td>Number of responses as a function of absolute error interval for</td>
<td>318</td>
</tr>
<tr>
<td></td>
<td>Subject C in the experimental group</td>
<td></td>
</tr>
<tr>
<td>Number of responses as a function of error interval for Subject D in the experimental group</td>
<td>319</td>
<td></td>
</tr>
<tr>
<td>Number of responses as a function of absolute error interval for Subject D in the experimental group</td>
<td>319</td>
<td></td>
</tr>
<tr>
<td>Number of responses as a function of error interval for Subject E in the experimental group</td>
<td>320</td>
<td></td>
</tr>
<tr>
<td>Number of responses as a function of absolute error interval for Subject E in the experimental group</td>
<td>320</td>
<td></td>
</tr>
<tr>
<td>Number of responses as a function of error interval for Subject F in the experimental group</td>
<td>321</td>
<td></td>
</tr>
<tr>
<td>Number of responses as a function of absolute error interval for Subject F in the experimental group</td>
<td>321</td>
<td></td>
</tr>
<tr>
<td>Number of responses as a function of error interval for Subject G in the experimental group</td>
<td>322</td>
<td></td>
</tr>
<tr>
<td>Number of responses as a function of absolute error interval for Subject G in the experimental group</td>
<td>322</td>
<td></td>
</tr>
<tr>
<td>Number of responses as a function of error interval for Subject H in the experimental group</td>
<td>323</td>
<td></td>
</tr>
<tr>
<td>Number of responses as a function of absolute error interval for Subject H in the experimental group</td>
<td>323</td>
<td></td>
</tr>
<tr>
<td>Number of responses as a function of error interval for Subject I in the experimental group</td>
<td>324</td>
<td></td>
</tr>
<tr>
<td>Number of responses as a function of absolute error interval for Subject I in the experimental group</td>
<td>324</td>
<td></td>
</tr>
<tr>
<td>Number of responses as a function of error interval for Subject J in the experimental group</td>
<td>325</td>
<td></td>
</tr>
<tr>
<td>Number of responses as a function of absolute error interval for Subject J in the experimental group</td>
<td>325</td>
<td></td>
</tr>
<tr>
<td>Number of responses as a function of error interval for Subject K in the experimental group</td>
<td>326</td>
<td></td>
</tr>
</tbody>
</table>
109. Number of responses as a function of absolute error interval for Subject K in the experimental group ......................................................... 326
110. Number of responses as a function of error interval for Subject L in the experimental group ................................................................. 327
111. Number of responses as a function of absolute error interval for Subject L in the experimental group ......................................................... 327
112. Number of responses as a function of error interval for Subject M in the experimental group ................................................................. 328
113. Number of responses as a function of absolute error interval for Subject M in the experimental group ......................................................... 328
114. Number of responses as a function of error interval for Subject N in the experimental group ................................................................. 329
115. Number of responses as a function of absolute error interval for Subject N in the experimental group ......................................................... 329
116. Number of responses as a function of error interval for Subject O in the experimental group ................................................................. 330
117. Number of responses as a function of absolute error interval for Subject O in the experimental group ......................................................... 330
118. Number of responses as a function of error interval for Subject P in the experimental group ................................................................. 331
119. Number of responses as a function of absolute error interval for Subject P in the experimental group ......................................................... 331
120. Number of responses as a function of error interval for Subject Q in the experimental group ................................................................. 332
121. Number of responses as a function of absolute error interval for Subject Q in the experimental group ......................................................... 332
122. Number of responses as a function of error interval for Subject R in the experimental group ................................................................. 333
123. Number of responses as a function of absolute error interval for Subject R in the experimental group ......................................................... 333
124. Number of responses as a function of error interval for Subject S in the experimental group ....................................................... 334

125. Number of responses as a function of absolute error interval for Subject S in the experimental group ........................................ 334

126. Number of responses as a function of error interval for Subject T in the experimental group ..................................................... 335

127. Number of responses as a function of absolute error interval for Subject T in the experimental group ........................................ 335

128. Number of responses as a function of error interval for Subject U in the experimental group .......................................................... 336

129. Number of responses as a function of absolute error interval for Subject U in the experimental group ........................................ 336

130. Number of responses as a function of error interval for Subject V in the experimental group ..................................................... 337

131. Number of responses as a function of absolute error interval for Subject V in the experimental group ........................................ 337

132. Number of responses as a function of error interval for Subject W in the experimental group ..................................................... 338

133. Number of responses as a function of absolute error interval for Subject W in the experimental group ........................................ 338

134. Number of responses as a function of error interval for Subject X in the experimental group ..................................................... 339

135. Number of responses as a function of absolute error interval for Subject X in the experimental group ........................................ 339

136. Number of responses as a function of error interval for Subject Y in the experimental group ..................................................... 340

137. Number of responses as a function of absolute error interval for Subject Y in the experimental group ........................................ 340

138. Number of responses as a function of error interval for Subject Z in the experimental group ..................................................... 341
139. Number of responses as a function of absolute error interval for Subject Z in the experimental group ............................................. 341
140. Number of responses as a function of error interval for Subject AA in the control group .......................................................... 342
141. Number of responses as a function of absolute error interval for Subject AA in the control group .......................................................... 342
142. Number of responses as a function of error interval for Subject BB in the control group .......................................................... 343
143. Number of responses as a function of absolute error interval for Subject BB in the control group .......................................................... 343
144. Number of responses as a function of error interval for Subject CC in the control group .......................................................... 344
145. Number of responses as a function of absolute error interval for Subject CC in the control group .......................................................... 344
146. Number of responses as a function of error interval for Subject DD in the control group .......................................................... 345
147. Number of responses as a function of absolute error interval for Subject DD in the control group .......................................................... 345
148. Number of responses as a function of error interval for Subject EE in the control group .......................................................... 346
149. Number of responses as a function of absolute error interval for Subject EE in the control group .......................................................... 346
150. Number of responses as a function of error interval for Subject FF in the control group .......................................................... 347
151. Number of responses as a function of absolute error interval for Subject FF in the control group .......................................................... 347
152. Number of responses as a function of error interval for Subject GG in the control group .......................................................... 348
153. Number of responses as a function of absolute error interval for Subject GG in the control group .......................................................... 348
154. Number of responses as a function of error interval for Subject HH in the control group ................................................................. 349

155. Number of responses as a function of absolute error interval for Subject HH in the control group ................................................................. 349

156. Number of responses as a function of error interval for Subject II in the control group ................................................................. 350

157. Number of responses as a function of absolute error interval for Subject II in the control group ................................................................. 350

158. Number of responses as a function of error interval for Subject JJ in the control group ................................................................. 351

159. Number of responses as a function of absolute error interval for Subject JJ in the control group ................................................................. 351

160. Number of responses as a function of error interval for Subject KK in the control group ................................................................. 352

161. Number of responses as a function of absolute error interval for Subject KK in the control group ................................................................. 352

162. Number of responses as a function of error interval for Subject LL in the control group ................................................................. 353

163. Number of responses as a function of absolute error interval for Subject LL in the control group ................................................................. 353

164. Number of responses as a function of error interval for Subject MM in the control group ................................................................. 354

165. Number of responses as a function of absolute error interval for Subject MM in the control group ................................................................. 354

166. Number of responses as a function of error interval for Subject NN in the control group ................................................................. 355

167. Number of responses as a function of absolute error interval for Subject NN in the control group ................................................................. 355

168. Number of responses as a function of error interval for Subject OO in the control group ................................................................. 356
169. Number of responses as a function of absolute error interval for Subject OO in the control group ........................................ 356

170. Number of responses as a function of error interval for Subject PP in the control group .................................................. 357

171. Number of responses as a function of absolute error interval for Subject PP in the control group .................................................. 357

172. Number of responses as a function of error interval for Subject QQ in the control group .................................................. 358

173. Number of responses as a function of absolute error interval for Subject QQ in the control group .................................................. 358

174. Number of responses as a function of error interval for Subject RR in the control group .................................................. 359

175. Number of responses as a function of absolute error interval for Subject RR in the control group .................................................. 359

176. Number of responses as a function of error interval for Subject SS in the control group .................................................. 360

177. Number of responses as a function of absolute error interval for Subject SS in the control group .................................................. 360

178. Number of responses as a function of error interval for Subject TT in the control group .................................................. 361

179. Number of responses as a function of absolute error interval for Subject TT in the control group .................................................. 361

180. Number of responses as a function of error interval for Subject UU in the control group .................................................. 362

181. Number of responses as a function of absolute error interval for Subject UU in the control group .................................................. 362

182. Number of responses as a function of error interval for Subject VV in the control group .................................................. 363

183. Number of responses as a function of absolute error interval for Subject VV in the control group .................................................. 363

xxv
184. Number of responses as a function of error interval for Subject WW in the control group ........................................... 364

185. Number of responses as a function of absolute error interval for Subject WW in the control group ........................................... 364

186. Number of responses as a function of error interval for Subject XX in the control group ........................................... 365

187. Number of responses as a function of absolute error interval for Subject XX in the control group ........................................... 365

188. Number of responses as a function of error interval for Subject YY in the control group ........................................... 366

189. Number of responses as a function of absolute error interval for Subject YY in the control group ........................................... 366

190. Number of responses as a function of error interval for Subject ZZ in the control group ........................................... 367

191. Number of responses as a function of absolute error interval for Subject ZZ in the control group ........................................... 367

192. Total practice minutes, total practice sessions, and highest difficulty level practiced as a function of each week of the training period for Subject A in the experimental group ................. 369

193. Total practice minutes, total practice sessions, and highest difficulty level practiced as a function of each week of the training period for Subject B in the experimental group ................. 370

194. Total practice minutes, total practice sessions, and highest difficulty level practiced as a function of each week of the training period for Subject C in the experimental group ................. 371

195. Total practice minutes, total practice sessions, and highest difficulty level practiced as a function of each week of the training period for Subject D in the experimental group ................. 372

196. Total practice minutes, total practice sessions, and highest difficulty level practiced as a function of each week of the training period for Subject E in the experimental group ................. 373

xxvi
197. Total practice minutes, total practice sessions, and highest difficulty level practiced as a function of each week of the training period for Subject F in the experimental group .......... 374

198. Total practice minutes, total practice sessions, and highest difficulty level practiced as a function of each week of the training period for Subject G in the experimental group .......... 375

199. Total practice minutes, total practice sessions, and highest difficulty level practiced as a function of each week of the training period for Subject H in the experimental group .......... 376

200. Total practice minutes, total practice sessions, and highest difficulty level practiced as a function of each week of the training period for Subject I in the experimental group .......... 377

201. Total practice minutes, total practice sessions, and highest difficulty level practiced as a function of each week of the training period for Subject J in the experimental group .......... 378

202. Total practice minutes, total practice sessions, and highest difficulty level practiced as a function of each week of the training period for Subject K in the experimental group .......... 379

203. Total practice minutes, total practice sessions, and highest difficulty level practiced as a function of each week of the training period for Subject L in the experimental group .......... 380

204. Total practice minutes, total practice sessions, and highest difficulty level practiced as a function of each week of the training period for Subject M in the experimental group .......... 381

205. Total practice minutes, total practice sessions, and highest difficulty level practiced as a function of each week of the training period for Subject N in the experimental group .......... 382

206. Total practice minutes, total practice sessions, and highest difficulty level practiced as a function of each week of the training period for Subject O in the experimental group .......... 383

207. Total practice minutes, total practice sessions, and highest difficulty level practiced as a function of each week of the training period for Subject P in the experimental group .......... 384
208. Total practice minutes, total practice sessions, and highest difficulty level practiced as a function of each week of the training period for Subject Q in the experimental group 385

209. Total practice minutes, total practice sessions, and highest difficulty level practiced as a function of each week of the training period for Subject R in the experimental group 386

210. Total practice minutes, total practice sessions, and highest difficulty level practiced as a function of each week of the training period for Subject S in the experimental group 387

211. Total practice minutes, total practice sessions, and highest difficulty level practiced as a function of each week of the training period for Subject T in the experimental group 388

212. Total practice minutes, total practice sessions, and highest difficulty level practiced as a function of each week of the training period for Subject U in the experimental group 389

213. Total practice minutes, total practice sessions, and highest difficulty level practiced as a function of each week of the training period for Subject V in the experimental group 390

214. Total practice minutes, total practice sessions, and highest difficulty level practiced as a function of each week of the training period for Subject W in the experimental group 391

215. Total practice minutes, total practice sessions, and highest difficulty level practiced as a function of each week of the training period for Subject X in the experimental group 392

216. Total practice minutes, total practice sessions, and highest difficulty level practiced as a function of each week of the training period for Subject Y in the experimental group 393

217. Total practice minutes, total practice sessions, and highest difficulty level practiced as a function of each week of the training period for Subject Z in the experimental group 394
CHAPTER I

INTRODUCTION

Background

Inquiry into the nature and origins of absolute pitch ability is by no means a recent phenomenon. Historically, the nature of this ability has been one of the most intriguing topics in the field of music perception, capturing the imagination of hundreds of writers. Scientific study of this ability began over a century ago, when Carl Stumpf (1883, 1890) investigated the pitch identification abilities of four subjects, himself included. Before the turn of the century, a tutorial existed that—it was claimed—would teach absolute pitch (Jadassohn, 1899), and the first experimental study dedicated to training and absolute pitch was published the same year (Meyer, 1899). Even today, most musicians seem to entertain a curious interest, as well as their own opinions, on the subject.

A substantial amount of literature has been written on nearly every aspect of the topic; unfortunately, much of it is of dubious merit. Large portions are either anecdotal or speculative, providing little substantive information that a scholar can rely upon. Moreover, portions of the literature that appear to have factual bases often lack the supporting data that could enable one to evaluate fully the merit of the researchers’ conclusions.
Definitions of Terms

Most musicians rely entirely on the use of relative pitch. People with good relative pitch are able to reproduce or recognize pitches by name, but only after being given a starting reference. In contrast, people with absolute pitch seem to carry around internal representations of pitches, so that they can reproduce or recognize them at any time.

Active absolute pitch will be used in this study to mean the ability to produce a specific tone without benefit of comparison to an objective reference. The other form of the ability, evidenced by the opposite form of the response task, is passive absolute pitch—defined here as the ability to identify a specific tone without benefit of comparison to an objective reference. The lack of an objective reference renders useless attempts to employ relative pitch. Passive absolute pitch is considered slightly more common than active absolute pitch, as well as less difficult to measure.

Description of the Problem

Absolute pitch, be it active or passive, is fairly rare and is often associated with an exceptional level of musicianship. Bachem (1955) reports that 1 in every 10,000 of the general population possesses absolute pitch, while Spender (1982) reports that 87% of a group of particularly gifted concert performers had the ability. Several early writers, notably Geza Révész (1913, 1954), Carl Seashore (1938, 1940), and Albert Bachem (1937, 1940) maintained that absolute pitch ability was not only rare, but was also a hereditary ability. Support for this position has dwindled over the years in the face of "the lack of any convincing supportive scientific evidence." (Ward & Burns, 1982, p. 435) Surprisingly, advocates of the hereditary viewpoint survive, as
witnessed by the recent work of Profita & Bidder.\textsuperscript{1} Figures presented by those who espouse the hereditarians' viewpoint (who also frequently identify themselves as possessors of absolute pitch) support the contention that absolute pitch ability is restricted to a smattering of remarkable individuals. Yet others, such as David L. Burge (1983), challenge this assertion. Like numerous previous writers who have advocated their absolute pitch tutorials, Burge contends that absolute pitch can be acquired by those who follow his training exercises.

References acknowledging the existence of Burge's method can be found in the current literature (Goldman, 1984; Irwin, 1984; Burge, 1985; Reubart, 1985; G. Mirken, 1987; M. Mirken, 1987; Steblin, 1987; Barkowsky, 1987), yet no writer offers any objective evidence as to the effectiveness of the tutorial. Steblin writes that "time will tell if Burge has indeed succeeded in this endeavor." (Steblin, 1987, p. 141) Furthermore, Barkowsky apparently takes exception to Burge's method:

\ldots which purportedly teaches the faculty of absolute pitch. Some of the terms used in Burge's eartraining course bear a deceptive resemblance to the chroma theory of absolute pitch, e.g., his term 'color hearing' parallels Bachem's idiom 'chroma apperception.' However, there is no relation between Burge's teaching and this writer's concept of absolute pitch. (Barkowsky, 1987, p. 42)

\textsuperscript{1}Stark (1985), Profita & Bidder (1988).
Statement of Purpose

The purposes of this dissertation were threefold. The primary goal was to investigate the effect of training on the acquisition of absolute pitch in adult musicians. Subjects were trained using Burge's method and materials, in a controlled experiment, with the hope that the evidence that it produced might help determine whether absolute pitch facility can be acquired by adults who were previously nonpossessors. The second goal of this study was to provide the reader with an extensive review of reports on previous attempts to train absolute pitch. In addition to providing factual information that could aid in the possible replication of important studies, the review included critical evaluations of methodology and data treatment. The third purpose of this study was to provide the reader with a comprehensive bibliography of literature on the topic of absolute pitch.

Need for the Study

At the outset of this study, there existed no comprehensive investigation of training for absolute pitch that was conducted with suitable scientific rigor. It was felt that if the results of this experiment established that the training method was effective, this would be an important study. As the first successful demonstration of teaching absolute pitch skills to a group of adults; it would support the tenuous findings of earlier training experiments; and produce major new evidence for a learning theory. If the treatment proved to be ineffective, the results would still be of great importance to individuals attempting to acquire absolute pitch skills, either through the use of the same training method, or through other methods, or on their own.
Limitations of the Study

This study focused entirely on the investigation of aspects of pitch judgment that are absolute, as opposed to relative, in nature. Furthermore, only identification tasks—involving passive absolute pitch as opposed to active absolute pitch—were explored. This is to say that subjects performed identification tasks; no production tasks were required.

Only isolated single tones were presented for identification, and only responses of pitch class name, not octave identification, were solicited. Identifications of simultaneities (dyads, triads, chords), keys, or of tonal centers of musical examples were beyond the scope of this investigation. Only piano tones were used for stimuli. No attempts were made to measure the identification of tones of any other timbre, and only tones of the standard 12-to-the-octave equal tempered (A = 440 Hz) scale were presented.

Subjects participating in this study were all college music majors. No investigations regarding either children or the elderly were made. No attempt was made to trace the heredity of absolute pitch among the families of any of the subjects. No correlations, or attempts to determine causation, were made between absolute pitch and other musical abilities, or between absolute pitch and any standardized tests or measures of musicality. No correlations of absolute pitch ability and any other factor (gender, race, age, handedness, etc.) were made.

Lastly, no study of the long-term retention of any absolute pitch skills that could potentially be acquired through the course of this experiment—beyond the endpoint of this experiment—was made.
CHAPTER II

REVIEW OF LITERATURE:

PREVIOUS ATTEMPTS AT TRAINING ABSOLUTE PITCH

Introduction

Since before the turn of the century, adults without absolute pitch have attempted to acquire this faculty through training. The research problem of attempting to instill absolute pitch skills reaches to the very heart of music education in this country and abroad. In fact, the two major schools of thought regarding sight-singing instruction (the fixed-do and movable-do systems of solfège) differ fundamentally on the basis of attention paid to the absolute identities of the musical pitches.

There are many good introductory essays on absolute pitch, but they invariably leave unanswered the question of whether absolute pitch is inborn or learned. This classic nature/nurture squabble does not persist because no one has bothered to examine the question; indeed, the published record of attempts to train absolute pitch extends back almost a century. Rather, the problem is that virtually all of these studies have been plagued by one methodological problem or another. Thirty published reports of attempts to develop absolute pitch have been reviewed; all imply that absolute pitch can indeed be learned.
Meyer (1899)

As far as can be determined, the first attempt to find experimental evidence that the memory for absolute pitch can be improved through training was made by Max F. Meyer (1873-1967). Meyer, who received his doctorate in 1896 from the University of Berlin, studied under the tutelage of Professors Carl Stumpf and Max Planck (Esper, 1966). Emigrating to the United States in 1898, Meyer spent the next year at Clark University before beginning a 30-year tenure at the University of Missouri. In 1899, he published the results of the pitch experiment (Meyer, 1899) that he and a colleague, Dr. Victor Heyfelder, had conducted in Berlin some 3 years earlier. Oddly enough, he and Heyfelder served as the only subjects in their experiment which lasted from March to October of 1895.

Meyer's report states that in the first 3 months they trained with tones produced by tuning forks, and in the last 5 months they trained with tones produced on the piano. In both instances the pitches of the tones were not identified by their pitch labels, but by their frequencies, which were available in tabular form. In the tuning fork portion of the experiment, Meyer and Heyfelder began with 9 tuning forks, and as training progressed, they increased the number of forks to a total of 16. As the number of forks increased, Meyer's accuracy decreased from 75% using 9, to 71% using 11, 67% using 14, and finally 59% using all 16. Likewise, Heyfelder's proficiency decreased from of 83% using 9 forks, to 78% using 11, 70% using 14, and finally 56% using all 16. The frequencies of the tuning forks were 100, 122, 150, 188, 220, 300, 400, 480, 680, 800, 960, 1200, 1600, 2400, 3200, and 4000 Hz. A display of these frequencies and their octave relationships (shown in Tables 1
and 2) demonstrates that even when all 16 forks were in use, there were only eight different pitch classes involved.

**Table 1.** Frequencies of tuning fork tones and the octave relationships between them in the Meyer and Heyfelder experiment.

<table>
<thead>
<tr>
<th></th>
<th>3200</th>
<th>1600</th>
<th>800</th>
<th>400</th>
<th>300</th>
<th>100</th>
<th>122</th>
<th>150</th>
<th>188</th>
<th>220</th>
<th>480</th>
<th>680</th>
<th>4000</th>
</tr>
</thead>
<tbody>
<tr>
<td>1600</td>
<td>2400</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>800</td>
<td>1200</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>400</td>
<td></td>
<td>300</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>122</td>
<td>150</td>
<td>188</td>
<td>220</td>
<td>480</td>
<td>680</td>
<td>4000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Table 2.** Frequencies of tuning fork tones in the Meyer and Heyfelder experiment reduced to one octave and shown in ascending order. Boldface numbers indicate the frequencies of the tones actually present.

<table>
<thead>
<tr>
<th></th>
<th>3200</th>
<th>1600</th>
<th>800</th>
<th>400</th>
<th>2000</th>
<th>2400</th>
<th>1000</th>
<th>1200</th>
<th>500</th>
<th>600</th>
<th>680</th>
<th>250</th>
<th>300</th>
<th>340</th>
<th>122</th>
<th>125</th>
<th>150</th>
<th>170</th>
<th>188</th>
</tr>
</thead>
<tbody>
<tr>
<td>1600</td>
<td>2400</td>
<td></td>
<td></td>
<td></td>
<td>1000</td>
<td>1200</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>800</td>
<td>960</td>
<td></td>
<td></td>
<td></td>
<td>600</td>
<td>680</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>400</td>
<td>480</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>200</td>
<td>220</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>240</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>110</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>120</td>
<td>122</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

For the portion of the experiment in which piano tones were used, Meyer stated that training started “with 10 pitches at intervals of a sixth. When the number of different pitches reached 20, the intervals were major thirds; when [the number of tones reached] 39, [the intervals became] whole tones.” (Meyer, 1899, p. 515) If this was the case, only 6 of the 12 pitch classes were used, due to the replication of the major third and whole tone at the octave, and the minor sixth at the double octave.

Unfortunately, only a few of the details regarding the training procedure are reported. From another source, however, we do know that Meyer and Heyfelder spent 1 to 2 hours per week practicing and that only the center of the keyboard was employed (Meyer, 1956). The initial levels of their
abilities to identify tones were not quantitatively reported before training began, but we know that—both being psychology students at the time—they regarded their abilities at memorizing tone height as quite poor. Neither experimenter was a practicing musician; in fact, neither had played the piano or an orchestral instrument since childhood. Notably lacking is any discussion of how the use of relative pitch was prevented during the testing. This omission makes it difficult to make any meaningful interpretation of their findings. Even more disturbing is their lack of discussion of the decision-making process used in judging the tones. We do know, nonetheless, that this experiment permitted each tone to be re-sounded as often as desired and that Meyer considered the faculty of absolute pitch largely as memorization of tone height, giving little regard to the concept of tone chroma. Meyer wrote sarcastically that “we never noticed any ‘tone chroma,’ but perhaps we just were too dull for that.” (Meyer, 1956, p. 719) Being nonmusicians, this was perhaps the case.

Regardless, by the time they completed the piano-tone portion of the experiment, Meyer had developed a terminal proficiency of 60% accuracy, and Heyfelder a 64% accuracy. Meyer’s statement that “errors surpassing the neighboring pitch on either side were quite rare” (Meyer, 1899, p. 516) is misleading; it is not clear if this was meant to indicate a semitone or not, inasmuch as the 39 test piano tones were separated by whole tones.

Meyer stated that at the conclusion of the experiment they felt that they had not reached their limits, but not being musicians, they concluded that the benefits of the newly acquired skill were not justified by the time required for its acquisition. Furthermore, at the time of the report (perhaps 3 years after
the 8-month training period) Meyer stated that “we have lost the greater part of what we had acquired, by the want of continued practice.” (p. 516)

Jadassohn (1899)

Das Tonbewußtsein: die Lehre vom musikalischen Hören (Jadassohn, 1899) was first published by German composer, conductor, and famed pedagogue Salomon Jadassohn (1831-1902), in the same year as the experimental report of Meyer.\(^1\) This musical textbook represented another in his highly acclaimed series of pedagogical works. Jadassohn studied piano with Franz Liszt, and theory and composition with Moritz Hauptmann. He received an honorary doctorate from the University of Leipzig and later became a professor at the Royal Conservatory of Music there.

Jadassohn believed that absolute pitch is gained by exercising one’s knowledge of relative pitch and can be attained through the systematic study and daily practice of the examples found in his book. Jadassohn advocated “acquiring the possession of a fundamental tone,” from which all other tones may be derived. He insisted that all training sessions be undertaken with an in-tune instrument, and recommended additionally the use of a tuning fork to help fix in mind the “fundamental tone.” He suggested the pitch “a,” but noted that the choice of “fundamental tone” depended on the student’s voice. Jadassohn strongly emphasized the need to begin each exercise by the testing one’s ability to remember the “fundamental tone.”

Jadassohn developed his techniques from his experience in teaching private piano lessons. He stated that by testing each of his students at the first

---

\(^1\)A second edition, translated into English, was posthumously published in 1905.
lesson, and by including ear-training in every succeeding lesson, he was astonished to find that many of his private piano students, after only a few minutes during each lesson, actually gained absolute pitch. He added that the use of his techniques in classes at the conservatory always produced satisfactory results, as well. The book contains 203 exercises, beginning with consonant intervals succeeded by dissonant intervals, dyads succeeded by triads, all followed by “dissonant” and “secondary” chords. This organization is typical of many ear-training texts, past and present. While there is little documentary evidence to support Jadassohn’s claimed success, there also is none to refute it.

Köhler (1915)

The German-American psychologist Wolfgang Köhler (1887-1967) investigated formal training for absolute pitch. Köhler received a Ph.D. in psychology in 1909 from the University of Berlin where he studied under Carl Stumpf, and he later was one of the founders of the Gestalt school of psychology. In 1915 he asserted that absolute pitch ability was not based on pitch memory, but on the memory of other tonal characteristics. He called these other features collectively the tonal body (Tonkörper) and described them as “brightness,” “vocality,” “volume,” and “intensity” (Köhler, 1915). Köhler believed that pitch (Tonhöhe) and tone body were independent, and that the quality of the tone body was recognized immediately, as opposed to involving the intervention of a stored memory image. As evidence for this belief he pointed out that individuals with absolute pitch who could label pitches immediately, were nevertheless frequently deficient in relative pitch abilities, were often confused by new timbres, were in many cases able to
identify chords better than single tones, and were prone to frequent errors at the octave, fifth, and fourth.

Köhler conducted an experiment in which he served both as the experimenter and the sole subject. He attempted to learn the seven pitch classes found in the C diatonic collection from C2 to B6 (USA Standard)—a total of 35 tones consisting of 7 of the 12 chromata in five different octaves. Köhler sought to redirect his attention away from the pitch (height) of the tone, focusing instead on the “body” of the tone. Initially, he found the training easy, and immediately eliminated all errors larger than an octave. After this first success, Köhler experienced some difficulty in making his judgments independently of pitch height, but thereafter his abilities again showed great improvement. Following only 2 weeks of systematic practice, he was able to identify correctly the pitches of 112 of 220 piano tones (51%). The majority of his errors were at the octave, fifth and fourth, a pattern which he believed corresponded to the usual error pattern of native possessors of absolute pitch. Such errors occur, stated Köhler, when attention is truly focused on the tone body, and not on the pitch. In fact, Köhler stated that, toward the end of the training period, his attention to the tone body was so fixed that he could ignore the pitch of the tone. Furthermore, he reported that two tones were apt to become confused if they possessed similar tone bodies and were reasonably close in pitch. Köhler found himself relying upon the attribute of “brightness” for the rough regional placement of a tone, and other characteristics unique to each tone for its specific placement. Some tones were “geschlossen (closed),” others were “zwiepaltig (having a double-meaning),” etc.
Although Köhler found that his judgments of absolute pitch were not impeded by the use of an unfamiliar piano, he nonetheless found that absolute pitch recognition of unfamiliar timbres was virtually impossible. The features that constitute the tone body are not present in pure tones, and Köhler likewise found their identification extremely difficult. The effect of Köhler’s training is difficult to ascertain, because he failed to document his level of ability prior to undertaking this training. Additionally, Köhler’s enhanced aptitude for pitch naming was unfortunately short-lived, for when he discontinued the training exercises, his increased facility gradually began to disappear.

Copp (1916)

Evelyn Fletcher Copp of Brookline, Massachusetts (dates unknown) expressed the view that the scientific community had placed far too great an emphasis on the hereditary basis of musical talent, and far too little on the learning aspect (Copp, 1916). Copp felt that if musical ability is hereditary, then “almost everyone possesses the heredity.” (p. 297) Arguing that the inheritance of a trait and the expression of a trait are two separate matters, Copp wrote that current studies of musical talent were faulty because they could not measure latent talent that had been inherited, but suppressed. Copp believed that while instances of great genius are sporadic and unaccountable, actual musical ability is much more common than most people suppose. She wrote that most musical ability is lost due to the lack of training during early childhood, when the mind and the ear are in their most sentient stage. Part of the evidence Copp used to support her ideas was her success in teaching absolute pitch to young children.
Copp based her statements on 20 years of teaching experience with hundreds of children, and added that the corroborative experience of the teachers she had trained would add hundreds of cases more. As direct evidence of her techniques, Copp's article included a photograph entitled "Acquiring a Supposedly Inborn Gift." It showed a young girl sounding tones on the piano, a young boy and girl blindfolded with their backs turned to the piano apparently naming the tones, and another young boy behind them locating the pitches on a large wooden musical staff. Copp (1916) wrote:

> We are only just beginning to learn what the normal ear is capable of, for instance in the matter of Positive Pitch, that is, ability to recognize and name musical tones. The lay public has been accustomed to consider Positive Pitch as a gift wrapped in the exclusive tissue of genius and doled out to the ultra musical only. One who can enter a room where a musician is singing or playing and say, "He is singing a high C, or baritone B," has hitherto been looked upon as a prodigy. This is by no means necessarily true. By proper training this power may be acquired, speaking very conservatively, by 80% of normal children. (p. 299)

Copp added that children who had been considered devoid of any musical talent, even the tone deaf, had, after a few months of training, been able to sing on demand middle C, to recognize it when it is played or sung, and soon thereafter to become equally familiar with the other musical tones. Copp acknowledged that the rate of progress varied, and that not every child could acquire absolute pitch by training, but she asserted that most could do so if they began at an early age.

Copp's etiology of absolute pitch combined aspects of both the hereditary and learning positions with the concept that learning must occur in early life if native ability is to survive.² While not speaking directly of

²An illustration of the importance of early learning and its relation to sensory perception is found in the work of Ackroyd, Humphrey, & Warrington (1974).
Copp’s work, Jeffress (1962) was the first to suggest that an early learning theory of absolute pitch might be closely related to the modern concept of “imprinting,” though this term came into vogue well after Copp’s work in 1916. Imprinting refers specifically to a form of early learning constituting the basis for attachment between a young animal and its parent(s). The main characteristic of imprinting is that it occurs during a critical period in development; if the behavior does not occur during a specific period in the animal’s life, it is unlikely to occur at all. It has been most clearly observed in species of birds that are able to walk or swim immediately after birth (Hess, 1958).

Anonymous (1916)

Evelyn Gough (Gough, 1922, pp. 15-16) reported a study done in 1916 by an unnamed student of Professor C. A. Ruckmick (1886-1961) at the Psychological Laboratory at the University of Illinois. In this study, training experiments were conducted using 2 “musical adults,” 2 children, and a class of ear-training students as subjects. Training began for the 2 adults and 2 children with the attempted memorization of C₄, E₄ and G₄, followed by the addition of all “white-key” tones in this octave. Later, the range was extended to two octaves. Next, the “black-key” tones were added. Conversely, the ear-training class, serving as a control group, studied tones from the entire keyboard in random order.

They found that a young blind woman who received a corneal implant at age 27 was unable to recognize simple visual patterns, even though the image-forming powers of her eyes were largely restored. She eventually reverted back to the life of a blind person.
After a training period of unspecified length, the size and the percentage of identification errors were reduced. It was concluded that accuracy varies according to the speed of judgment, since the average time for correct responses was shorter than for incorrect ones. The experimenter also noted the tendency for the children to make all judgments relative to another note, usually C.

Maryon (1919)

Another teaching method introduced by a famed musician was that of the English composer Edward Maryon (1867-1954). Maryon’s early training at the Royal Academy of Music included piano study with Oscar Beringer, and harmony, counterpoint and composition studies with Ebenezer Prout and Sir George Macfarren (Hipsher, 1934). Maryon later went to Paris to specialize in the interpretation of Chopin’s music with I. Libich, one of Chopin’s pupils. Maryon’s opera L’Odalisque won the Grand Prix and Gold Medal from the French Republic in 1890. Shortly thereafter he was awarded an honorary doctorate and was elected to membership in several prestigious organizations. Maryon was also a student of archaic languages, physics, and philosophy, studying most notably with Dr. Carl Hänsen, the father of modern hypnotism. Maryon first visited the United States in 1892. From 1914 to 1919 he resided in Montclair, New Jersey, where he established a conservatory supported by an exchange program for English and American music students. There he developed and codified his theories correlating color and sound in
the teaching of music. These theories were published under the name *Marcotone*.\(^3\)

Maryon’s techniques were first announced in 1916 (Kramer, 1916), before he formally introduced them in New York in 1917. The standard edition of his book, *Marcotone*, was published in 1919.\(^4\) Maryon held that because the source of both light and sound is vibration, the two co-exist in a natural relationship. Visual color, registered through the eye and received by the brain is, he asserted, judged according to its just measurement as light waves. Light is based on the law of atavism, or the transmission of light waves through the “element” ether. Tone is registered through the ear and received by the brain, but not measured by the brain according to judgments as sound waves. Instead, through the process of color-thought, specific ratios of color can be correlated to specific ratios of tone. In other words, if a color is mentally retained, the brain will signal this fact with a response of the exact correct tonal frequency from the mouth.

The book contains charts indicating the relationship between each tone in the octave above middle C and its corresponding color. The instructor of the course, if available, would use a color keyboard, a Marcotonograph (a Marcotone pitch-pipe which produced “tone-colors” as well), and other Marcotone appliances. The Marcotone scale was taught in a series of 12 lessons in which the 12 fundamental tones are divided into six primes and six complements. The six primary tone-color names and their diatonic spellings

---

\(^3\)Maryon stated that he derived the term Marcotone, a hybrid acronym, from Sanskrit, with “MA” meaning to measure, “R” (raga) meaning by scaling, “CO” meaning color, and TONE.

\(^4\)A condensed “practical” version, which omitted much of the theoretical explanation of Marcotone as a science, was released in 1924.
are Red = C, Orange = D, Yellow = E, Green = F#/G♭, Blue = G#/A♭, and Violet = A#/B♭. These follow the accepted order of the components of white light that are found, for instance, in a rainbow, or from a prism.⁵ All tones and colors, given in their order to be learned, are shown in Table 3.

Table 3. All colors (shown with color names and frequencies in Angstroms) and tones (shown with pitch labels and frequencies in Hertz) in the order in which the Marcotone system states they are to be learned.

<table>
<thead>
<tr>
<th>Lesson</th>
<th>Tone Color</th>
<th>Color</th>
<th>Pitch</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Red</td>
<td>6870</td>
<td>C</td>
<td>256</td>
</tr>
<tr>
<td>2</td>
<td>Yellow</td>
<td>5601</td>
<td>E</td>
<td>322.5</td>
</tr>
<tr>
<td>3</td>
<td>Blue</td>
<td>4555</td>
<td>G#/A♭</td>
<td>406.5</td>
</tr>
<tr>
<td>4</td>
<td>Orange</td>
<td>6164</td>
<td>D</td>
<td>287.5</td>
</tr>
<tr>
<td>5</td>
<td>Green</td>
<td>4919</td>
<td>F#/G♭</td>
<td>362</td>
</tr>
<tr>
<td>6</td>
<td>Violet</td>
<td>4104</td>
<td>A#/B♭</td>
<td>456.5</td>
</tr>
<tr>
<td>7</td>
<td>Orange-Red</td>
<td>6472</td>
<td>C#/D♭</td>
<td>271.5</td>
</tr>
<tr>
<td>8</td>
<td>Orange-Yellow</td>
<td>5865</td>
<td>D#/E♭</td>
<td>305</td>
</tr>
<tr>
<td>9</td>
<td>Yellow-Green</td>
<td>5233</td>
<td>F</td>
<td>342</td>
</tr>
<tr>
<td>10</td>
<td>Green-Blue</td>
<td>4737</td>
<td>G</td>
<td>384</td>
</tr>
<tr>
<td>11</td>
<td>Violet-Blue</td>
<td>4241</td>
<td>A</td>
<td>430.5</td>
</tr>
<tr>
<td>12</td>
<td>Violet-Red</td>
<td>3976</td>
<td>B</td>
<td>483.5</td>
</tr>
</tbody>
</table>

The 1919 edition contains 23 progressive sight-singing examples which must be sung at the correct pitch (no transposition allowed). Maryon (1924) had such confidence in his technique that he predicted:

Once this new and vital factor in evolution is realized by those responsible for the nation’s education, we shall become a race of natural musicians. Song will be a common, everyday gift, even as speech is today; a new joy will have come into the hearts and minds of the people and a new and more harmonious epoch of life will fill the earth. (p. 18)

---

⁵In order, ranging from longest wavelength to shortest wavelength: red, orange, yellow, green, blue, indigo, and violet.
Although there is no experimental evidence verifying the merits of the Marcotone system, Maryon (1919) wrote that “Fifty musicians in the U.S.A. have studied and mastered the Science of Tone-Color, and are prepared to assist the inventor of ‘Marcotone,’ as demonstrators, lecturers and teachers.” (p. 7) Furthermore, Kramer (1916) reported that Maryon taught his system in Montclair, New Jersey, with complete success.  

Gough (1922)

Gough conducted her own studies on the effect of extensive practice on absolute pitch in 1917 and 1918, using female undergraduates at Smith College (Gough, 1922). She ran nine different experiments, the most significant involving the identification of piano tones by three separate groups of subjects. The first group contained 9 advanced psychology students and 2 instructors, all of whom participated for the majority of the school-year. They were given a pretest involving all 88 piano tones presented in a quasi-random order. Seated with their backs to a new Knabe upright piano tuned to A₄ = 445 Hz, they faced a cardboard facsimile of a piano keyboard with octave designations printed on it. Tones, presented at a medium volume, were repeated as often as the subjects desired. The required verbal identifications included octave designations, and response times were recorded with a stopwatch. Then, while subjects individually practiced naming piano tones

---

6Other early attempts to train chromaesthesia include the experiments of Kelly (1934) and Howells (1944). Howells’ work is especially interesting not only because the college student-employees serving as subjects were reported to the faculty as being unreliable or inefficient and were terminated following poor experimental performance, but because the organ tones of C₄ (261 Hz) and G₄ (392 Hz) were presented with a red (carmine) and a bluish-green color, respectively. These selections exactly follow the Marcotone relationship between color and pitch.
in four 10-minute sessions per week, they were tested once each week on all 88 piano tones, as in the pretest. Six members of this group used the entire keyboard from the beginning, while four members were instructed to practice in the middle octave at first and to add octaves one at a time, first above, then below, at 2-week intervals. Because of disappointing results, the second technique was abandoned after several months. During testing, these participants had no restrictions on the duration of the stimulus tones, or the number of repetitions, or on their response time.

The second group consisted of 80 students (mostly college juniors), who were taking a required undergraduate psychology course; they participated for the 6 weeks between spring vacation and the end of the school year. Participants in this group practiced for two 10-minute periods, then took a weekly test involving all 88 piano tones, a test similar to the one used with the first group. Because of the larger group size, testing was given in groups of 20 instead of individually, and responses were written. Subjects still faced a keyboard diagram with the piano behind them, and their responses included pitch and octave designations. Because there was no longer a way to record response times individually, participants were given 10 seconds to hear several presentations of each tone, and 10 seconds of silence in which to respond.

The third group comprised 3 subjects with exceptional musical ability. Two were students (a junior and sophomore) who apparently already possessed some degree of absolute pitch ability; one was the instructor of ear-

---

7A major drawback to such individual training is the probable lack of the ability to respond to pitch memory cues without already knowing that the piano tone about to be sounded will be higher or lower than the last one. The subject will also already know that it will be a white or black key by the feel.
training and harmony at the Department of Music, who did not. Participants in this group followed a schedule similar to that of the second group.

All subjects were tested on their abilities to identify correctly the pitches of piano tones, including correct octave designations. As for training techniques, the subjects in the first group were instructed to devise their own methods for retaining the pitches of the different piano tones without reference to outside criteria. Practicing alone, these subjects were tested individually by the experimenter, or by each other. Because these subjects freely and independently created their own training techniques, it is not known what method or methods, if any, the data gathered by Gough actually represent. There is no indication that the subjects in the first group were ever informed of the outcome of their weekly tests, so these tests were probably of little or no benefit to them in acquiring absolute pitch. In contrast, the members of the second group were tested in groups and informed. Thus, the results of their weekly tests may have benefited them.

Unfortunately, Gough reported most of the quantitative data regarding subjects' progress through the use of group means. Although the frequency distributions of correct responses over the total number of weekly tests indicate correct responses for the first, second, and third groups as 17.7%, 8.7%, and 40.9% respectively, all data concerning improvement from the first test to the last are given as a combination of the three. Thus, distinct group means cannot be obtained. Because Gough gave individual practice curves for only 5 of the subjects, the vast majority of individual performance characteristics is lost. All that can be determined is that among the 5, accuracy ranged from one subject who could correctly identify 83 of the possible 88 tones, to another who could correctly identify only 1. While such findings are
of dubious value, due to differing lengths of practice time, differences in
testing procedure, and distinct differences in skill level at the outset, Gough,
nonetheless, reported that the mean percentage of correct responses from the
first test to the last increased from 8.5% to 13.1%, while the mean error
decreased from 5.4 to 4.0 semitones.

Mull (1925)

In 1925, Helen Mull published an account of three separate training
experiments conducted at the Harvard Psychological Laboratory. These
experiments investigated the role of attention in the acquisition of absolute
pitch (Mull, 1925). Mull used a total of 12 subjects; in addition to herself, she
used 7 graduate psychology students at Harvard University (6 males and 1
female), 2 undergraduate senior females also majoring in psychology, and 2
subjects who held doctoral degrees (1 female from Clark University, the other
from Harvard University, gender not specified). Three of the 12 subjects
lacked formal musical training and apparently experienced difficulty even in
vocally reproducing a tone well within their vocal ranges. Two of the
subjects had taken a few piano lessons, but no longer continued to play. The
remaining 7 subjects had received musical training, and still engaged in some
form of musical activity at the time of the study. According to pretest results,
none had absolute pitch. Their levels of musicality were quantified,
however, by the use of Seashore’s *Measures of Musical Talents* (1919).

Mull’s first experiment sought to determine if “absolute pitch” could be
obtained for one pitch only.\(^8\) Mull assumed that progress could be accelerated

\(^8\)This would not meet the criterion of absolute pitch as the term is commonly
used.
by limiting the number of pitches, compared to a similarly reduced exposure to all 12 pitches. By extension, she also assumed that if this faculty could be acquired for one pitch, it could be learned for all the pitches, given enough training. A group of 6 subjects, including both genders and representing a mixture of educational and musical ability levels, were pretested using single, "diatonic scale" tones from within "middle" three octaves of the pitch region. All tones used during training and ensuing testing sessions were produced either on an Appunn tonometer or an Ellis harmonical. The group was divided into two groups of 3 to determine if a stimulus tone, even though not directly attended to, would be of benefit in acquiring absolute pitch.

The two groups trained according to the following regimens. Three subjects trained an hour per week for most of the academic year. They were presented a middle C (C4 = 264 Hz) of 5-second duration once a minute for 15 minutes. They were instructed to attend to the tone during its presentation; they were allowed, in addition, to sing it, or do anything else that might aid them while they were observing the tone. For the next 55 seconds, however, they were told to distract themselves by reading or studying. The other 3 subjects were given the same instructions regarding their periods of attention and inattention, but, in their case, the tone was sounded continuously for all 15 minutes. These 3 subjects trained for 1 hour per week during the first semester, and 2 of the 3 trained an additional hour per week (2 hours total) during the second. At the conclusion of each 15-minute training period, both groups provided introspective comments, and following a brief period of auditory distraction, the testing began.

Subjects were presented a series of tones "chosen from among the naturals" from the two-octave range from C3 to B4, and, following a two-
alternative forced-choice model, were asked to respond whether the tone was or was not C4. Upon a "yes" response feedback was given, including the direction and magnitude of the error if necessary and the reinforcement of C4 by sounding it for 5 seconds. Tests were usually conducted 20 times each session, each isolated by a short auditory distraction.

After 1 month, in the second phase of the experiment, the test stimuli used were much closer in pitch than previously. The tones now had frequencies of 232, 240, 248, 256, 264, 272, 280, 288, and 296 Hz. These frequencies represent four 8-Hz increments above and below middle C, the tone introduced in the previous training exercises. The subjects were still instructed to indicate whether or not the tone they heard was middle C, but now no correction was given and the series no longer ended after a "yes" response. Moreover, the subjects were told that middle C might be included more than once, or not at all. For the sake of ease of computation, all nine tones were actually presented for judgment. Subsequently, both phases were alternated at 2-week intervals.

For the first portion of the experiment, regarding confusion with the other members of the C major scale, Mull reported that prior to training the average percentage of responses to middle C for all subjects was correct 40.4% of the time, and that the average error of all responses judged as middle C was 2.85 semitones. After training, the average accuracy of responses rose to 82%, while the average error fell to 0.33 semitones. She found that even though both groups attended to the stimuli for the same duration of time, no striking superiority was displayed by the group that received the continuous stimuli. She also reported that most of the improvement took place initially and that
a high degree of transfer took place when tones produced on various organ stops replaced the test stimuli produced on laboratory instruments.

Though impressive, Mull's averages are suspect. A careful inspection of the individual subject records shows that subjects rarely, if ever, received the same number of exposures to tones in the pretest. In fact, 2 of the 6 subjects in this experiment heard no occurrences of middle C at all. Moreover, no pretest data whatsoever were reported for a male psychology student who placed in the 3rd quartile of the Seashore test, because he did not know the note names prior to the commencement of the training sessions. The procedures caused other problems. Even though stimuli were separated by auditory distractions, the subjects received correction at some point during each test. Furthermore, their task was not to identify the name of the tone, as one with absolute pitch would normally be expected to do; instead, subjects needed to respond only whether the tone presented was or was not middle C. Nonetheless, recomputations of Mull's data for the 5 subjects for which complete records are given, when combined, yield an impressive increase from 52% to 89% of correct responses to all of the stimuli in the C diatonic collection (not just middle C) after training.

For the portion of the experiment dealing with tones separated by 8-Hz intervals, Mull reported that from the first test to the last, the average percentage of correct responses increased from 43% to 57%, while the average error was reduced from 0.68 semitones to 0.29 semitones.9

The quickness of learning, the disparity of results among subjects, and the marginal effect of the constant stimulus led Mull to focus on the variable

---

9This degree of error is not appreciably greater than the 0.33 semitone error that occurred when the stimulus tones were separated by full semitones.
of subject attention. In her second experiment, Mull sought to direct all of the subjects' attention to the tone in isolation, to the exclusion of any possible harmonic or melodic significance. She hoped that this narrowed task (Aufgabe) would positively influence 4 new subjects' success in acquiring absolute pitch. These subjects included 3 male psychology graduate students and a female college senior in psychology. They were given the same pretest used in the first experiment and, in addition, similar pretesting on the tone B₃ (248 Hz)—again making decisions as to whether or not the tone presented was B₃. A single training session presented the tone B₃ for 5 seconds, once every minute for 10 minutes. Testing to determine the retention of this tone consisted of a presentation in quasi-random order of all of the notes in the C diatonic collection from F₃ to E₄ inclusive, and also A♭₄ and B♭₄. The added tones extend beyond the range of the others and apparently were added "to offset the possibility of melodic cues."¹⁰ The subject was asked only to decide if each tone presented was or was not B₃. These nine-tone strings were constructed so as to minimize possible melodic interpretation. Tests additionally measured reaction times to determine the relationship between the speed and accuracy of judgments. The testing followed a graduated schedule; they followed 3 minutes, 1 hour, 1 day, 1 week, and 1 month after the training session. Mull reported that the heightened attention resulting from the narrowed focus produced positive results in 3 of the 4 subjects, and that response times for correct responses were slightly shorter than for incorrect responses (means of 5.9 and 6.0 seconds, respectively). Regarding the rate of forgetting, Mull reported that learning did indeed persist for a month. It is certain, however, that the tests themselves were conducive to

¹⁰Ibid., 483. Never did a subject identify either of these two tones as B₃.
subject learning; the average percentage of correct judgments generally got better, not worse, as the times following the training period got longer! The seven C major diatonic test tones produce a highly meaningful context, recognizable even to nonmusicians. In this context, Mull could not have chosen a more obvious tone to be identified than B, which happens to be the leading tone in this collection. Also, this tone forms a tritone with the lower boundary tone of the set, thereby producing a rare and easily identifiable interval.

The response task of Mull’s third and final experiment bears the greatest similarity to what one normally expects in absolute pitch behavior. Two subjects participated in the experiment: subject P, a holder of a doctorate from Clark University, and Mull herself, who was a graduate student at Harvard University at the time. Both subjects had received enough formal music training to be currently participating in some form of musical activity, and both had placed in the 1st quartile of Seashore’s *Measures of Musical Talents*. They set out to learn all of the chromatic tones from G3 (196 Hz) to F#4 (369 Hz) inclusive. Subject P was given the same pretest used in the previous experiments and trained using the Appunn tonometer. Training consisted of presentations of each of these 12 tones in the middle range of the piano. After training was well underway, and in alternate sessions, subject P was asked to give her response immediately, as opposed to a naturally occurring response. Training is described only as being intensive, yet this does not necessarily mean that training was extensive: no details regarding the length or total number of sessions is provided. Tests, consisting of two presentations of each tone in an irregular order, were given three times each

[^1]: Ibid., 487.
week for 4 months. In these tests, auditory distraction was introduced to prevent the use of relative pitch. In addition, Mull trained herself to these same 12 pitches in the summer of 1924 using common brass organ reeds randomly selected at large time intervals. No pretest data were recorded for Mull inasmuch as she had acted as the experimenter in the first experiment.

The scores of subject P increased from 7% correct judgments with an average error of 4.12 semitones on the pretest, to 38% correct judgments with an average error of 1.50 semitones on an intermediate test, to 62% correct judgments with an average error of 0.65 semitones on the final test. Only final test scores are given for Mull and they indicated 75% correct responses with an average error of 0.27 semitones. Mull additionally reported that learning was largely immediate, and even more so than in the two previous experiments. However, in contrast with the previous experiments, the responses of subject P became less accurate as her response time decreased. Of course, no response times are available for Mull because she trained alone.

Mull concluded that the mechanism used in the judgment of absolute pitch was largely based on timbre. She reached this conclusion for several reasons. While kinesthetic aids such as singing or whistling were used by some subjects, their use gradually diminished, even though they were occasionally called upon when subjects were faced with a particularly difficult decision. Even attention to pitch proved less fruitful than attention to timbre, although both were often used in combination. She asserted that, typically, pitch was used first for initial rough judgments and timbre was used for finer judgments. She wisely qualified her assessment, and discounted the role of chroma, with this statement:
There is no objective evidence which necessitates the postulating of an attribute of quality, or tonality, which recurs in successive octaves, and upon which all judgments of absolute pitch rest; nor do the introspective reports disclose the existence of any such attribute. (p. 492)

But, even this qualification seems inadequate. Subjects in the first experiment were tested only with stimuli from the C diatonic collection ranging from C₃ to B₄, and were asked only whether the tone sounded was or was not C₄. The chance of making an incorrect answer indicating an octave error was at most 1 in 14 since the confusion would also have to involve a misjudgment of the lower boundary tone. The subjects in the second experiment were instructed to identify the tone B₃ and were tested using stimuli from the C diatonic collection ranging from F₃ to E₄, with A₄ and B♭₄ added. This set affords no possibility of committing an octave error. In the third experiment involving the absolute identification of the chromatic collection from G₃ to F♯₄, neither subject made an octave error, and this is not surprising. The lack of opportunity for octave errors is an artifact of the testing procedure itself.

**Mal'tseva (1925)**

At the National Conservatory in Moscow, U.S.S.R., Esfir' Abramovana Mal'tseva also experimented with the acquisition of absolute pitch (Mal'tseva, 1925). Mal'tseva used 5 women as her subjects: 1 was a student of vocal pedagogy, 1 was enrolled in a technical music course, and the other 3 were characterized as “quite musical” although none of them specialized in music. Sessions with the subjects took place weekly, occasionally every 5 days. Mal'tseva attempted to teach her subjects to identify the names of 36 tones from all 12 pitch classes in the three middle octaves of the piano.
Subjects were instructed to form mental images of the tones they were about to judge but were not permitted to vocalize them. The training sought to focus the subjects' attention systematically on the timbre rather than on the pitch height of piano tones, a goal nearly identical to Köhler's. Mal'tseva rather unsystematically allowed her subjects to attend various numbers of training sessions, and used no standardized method for reporting their progress. She provided no pretest measure, other than the score recorded after the first training session. The following descriptions, translated by Teplov (1966), comprise all the information that is available.

Subject P, one of the nonmusicians, attended 12 training sessions. Her mean percentage of correct responses was 11% for the collective scores of the first three sessions, and 25% for those of the last three sessions.

Subject S, another nonmusician, attended 22 training sessions. Her mean percentage of correct responses was 18% for the collective scores of the first six sessions, and 36% for those of the last six sessions.

Subject K, the last of the nonmusicians, attended 24 training sessions. Her mean percentage of correct responses given in the early sessions is not given, and the only quantitative data Teplov recounts are that the mean percentage of correct responses during her last sessions varied between 30% and 35%.

Subject B was the student enrolled in a technical music course and she attended 25 training sessions. Her mean percentage of correct responses was 19% for the collective scores of the first five sessions, and 44% for those of the last five sessions.

Subject V was the vocal pedagogy student, and she attended nine training sessions. Her mean percentage of correct responses for each of the
nine sessions was 26%, 43%, 54%, 48%, 54%, 46%, 43%, 76%, and 67% respectively.

**Wedell (1934)**

Claiming a change in posture from the older absolute pitch literature that portrays subjects learning to respond to individual tones, Carl H. Wedell (1934) of the Psychological Laboratory at Princeton University expanded upon the previous work of Gough (1922) and Mull (1925). Wedell suspected that even in the absolute judgment of tones, subjects still base their judgments in part on pitch relations. He sought to ascertain whether persons can acquire absolute pitch, and if so, to determine the limits of their accuracy. Moreover, he sought to determine, through the use of different series of tones, whether subjects learn to recognize individual tones, or, instead, learn the extent and character of a whole scale.

Wedell used “relatively pure tones” for stimuli in his experiment. They were generated by a Western Electric vacuum tube oscillator, and although he could not accurately measure their harmonic content, he estimated their upper partial strength to be no greater than that of tuning fork tones. These tones were presented through headphones to subjects who were individually tested in a soundproof room; communication with the experimenter, located outside, occurred only by telephone. The stimulus tones were apparently presented monaurally. Wedell stated that “both E [xperimenter] and O [observer] wore two headphones, one in which they heard the stimulus tone, and one through which they received communication from each other.” (Wedell, 1934, pp. 488-89) With the aid of
a large chart placed before them, subjects verbally identified the frequencies of
the stimulus tones in the series they were learning.

Wedell experimented with 7 subjects, none of whom was currently
active in the field of music. Four of the subjects had never studied music; the
others had studied piano only in their childhood. Three of the subjects were
graduate students and 4 were departmental staff members.

Wedell’s first experiment required 4 subjects, including the 3 who had
never studied music, to identify 25 tones ranging in frequency from 50 Hz to
7400 Hz, a range roughly equivalent to that of a standard piano, but one
octave higher. Tones were all presented at an equal loudness of 40 dB above
the minimum threshold of perception, using a graph of equal-loudness
contours.\(^{12}\) Tones were separated by approximately equal numbers of
difference limens (dl’s). Because he lacked an equation for the curve of the
size of the dl versus frequency, and because the size of the dl varies
continuously, Wedell was forced to space his tones by the “method of
approximation.” Wedell’s experiment thus was the third to involve the use
of stimulus categories not corresponding to musical intervals, following both
Meyer and Heyfelder’s tuning fork experiments, and Gough’s tones at 8 Hz
increments about middle C. His 25 tones were separated by an average of 55.4
dl’s, which corresponds to a range of a musical interval slightly larger than a
minor sixth between the two lowest tones, to an interval slightly larger than a
minor third between the two highest tones.

\(^{12}\)A series of controlled 10 dB variations in intensity occurred during three
consecutive sessions at a point roughly 80% through the training. These
produced no variation in the average error for the four subjects. Wedell
concluded that they were not using loudness as a cue in their judgments
anyway.
Experimental sessions took place three times a week for slightly over 7 weeks; 2 subjects completed 23 sessions, the other 2, 22 and 21. Each session lasted less than an hour, and each began with a trial repetition of the series. During testing, a tone was presented for 5 seconds, an immediate identification was made, and then approximately 10 seconds elapsed before the next presentation. No correction was given. This procedure was repeated until all 25 tones were presented. Testing was followed by a learning period consisting of three presentations of the series with corrections provided after each mistake. In both testing and training, singing by the subjects was permitted, although they were encouraged to respond quickly.

Tones in the series were arranged in a quasi-random order to help prevent the subjects from simply learning orders rather than individual tones. Presentations contained an equal representation of small and large intervals, with tones adjacent in the series never following one another. The order, however, usually remained fixed for each of these sessions, although the starting point varied each time.\textsuperscript{13} In the penultimate session the frequencies of 9 of the 25 tones were modified, unannounced, from 1/3 to as much as 2 1/3 semitones, and in the final session the frequencies of all 25 of the tones were raised or lowered from 1/5 to as much as 4 full semitones.

Inspection of the individual subjects’ learning curves shows a sharp reduction in the average error in dI’s from the first to the fourth session; by the eighth session, levels of performance remained fairly constant. The sharp

\textsuperscript{13}Ward (1963b, p. 36) has suggested that memorization might have played some part in this study, but Wedell’s original report notes that the presentation order was changed once for each subject at a point roughly 70% through their training to determine if the order of the tones and not the tones themselves had been learned. This new order had little or no effect on performance, indicating that the subjects had indeed learned all 25 tones.
improvement early in training, followed by a plateau with moderate additional improvement is typical of most learning curves. The average error was about 125 dl’s (2.3 categories) before training, and was quickly reduced to about 75 dl’s (1.4 categories) by the fourth session. After a brief period of variation, the eighth session established an average error level of about 59 dl’s (1.1 categories) which remained fairly constant for each subject for the rest of the experiment. This represents an average error of approximately 3 semitones. The modification of the frequencies of some, and then all of the tones in the last two sessions produced errors that were not very different from those of previous sessions; in fact, 2 of the subjects did not even notice the alterations. Wedell found that, contrary to the distribution of absolute pitch errors reported in most previous studies, his subjects were poorest in the middle of the stimulus range where the tones fall between 500 Hz to 1910 Hz.

The 4 subjects were then given an extended layoff from training. The break lasted 2 months for 3 of the subjects, and 4 months for the other. When the subjects returned, the retraining of the original 25 tones required two sessions for 3 of them and three sessions for the other. Following the relearning of the original tones, subjects were given two test presentations, 48 hours apart, of a new series comprising the original 25 tones and 24 new ones, whose frequencies lay midway between the originals. This addition raised the total number of tones in the series to 49. Although the subjects had not received training with this new series, the accuracy of 3 of the 4 subjects’ responses were, surprisingly, not greatly reduced by the more compact series. In fact, the average error for all 4 subjects rose only to approximately 72 dl’s (1.3 of the original categories). Wedell also found that subjects made about
the same number of errors in judging the previously learned tones in the combined series, as they did in judging those that had been added. He concluded that training had the effect of establishing a familiarity within a frequency range without establishing absolute responses to specific tones. Wedell designed a second experiment to replicate and test the generalizability of these findings.

Wedell’s second experiment used 3 different nonmusicians as subjects, although 2 of them had received piano training (2 and 6 years, respectively). These subjects first learned a series of 5, then 9, then 13, and finally 17 different stimulus tones, again separated by approximately equal numbers of dl’s. These new divisions of roughly the same frequency range used in the earlier experiment represent average separations of 333.7, 167.4, 111.2, and 83.8 dl’s, respectively. Training sessions still began with a test session, but the number of series presentations was reduced from three to two. Although the series were not considered to be learned until the subject obtained one errorless test repetition, Wedell stated that some subjects were advanced to the 13- and 17-tone series before this requirement was met. While this recourse saved some time, it was permitted only when a subject was no longer making progress on the current series. At the completion of training, subjects were given two test repetitions, 48 hours apart, of the 25-tone series used in the earlier experiment. This series was, of course, novel to these subjects.

All 3 of Wedell’s subjects were quickly able to learn the five tones separated by an average of 333.7 dl’s. This separation is roughly equivalent to two octaves and a minor seventh between the two lowest tones, and diminishes continuously to about a perfect eleventh between the highest two.
Two subjects required only two sessions to acquire mastery of this series, while the other required three. The nine-tone series did present more of a problem to the subjects, but nevertheless, one subject mastered it in five sessions, the others in seven and eight. In this series the 167.4 dl average separation between tones is roughly equivalent to a major thirteenth between the two lowest tones, and diminishes steadily to slightly less than a major sixth between the highest two.

None of Wedell's subjects, however, could master the series of 13 tones separated by an average of 111.2 dl's, and all had conceded as much after 7, 10 and 13 sessions, respectively. The two lowest tones were now separated by about a minor tenth, while the two highest were separated by less than a tritone. In spite of their failure to master the 13-tone series, the subjects began to train with the 17-tone series and their performance again ended in failure. One subject trained for only two sessions, the others for three, but surprisingly the subjects' response accuracy was not appreciably different than it had been for the 13-tone series.

The subjects were next given two test repetitions of the series of 25 tones separated by an average of 55.4 dl's. One subject showed dramatic improvement over his performance with the 17-tone series, while the other 2 stayed at about the same level. The average error on this novel series ranged from 12 to 70 dl's, with median error of about 50 dl's. This represents an average error of approximately 2 1/2 semitones. These results are quite similar to the 59 dl, 3-semitone error obtained by the training method used in the earlier experiment.

To summarize, Wedell sought to calculate the average separation of tones in each series in semitones. To do so, he divided the average separation
in dI's for the series on which the 3 subjects had trained (which actually varied from each pair of tones to the next) by 20, the number of dI's he felt were contained in a semitone throughout most of the "tonal scale."

Compounding these errors, he concluded that:

Unmusical observers can learn accurately and easily to recognize tones that are eight and one third semitones apart [9-tone series], but they fail to learn to judge the tones correctly when the interval is decreased to five and one half semitones or less [the 13-tone series]. (p. 503)

Wedell also concluded that:

The person who acquires a pitch-naming ability is really acquiring a knowledge of an auditory extent, rather than of individual notes within that extent. Otherwise the errors should have increased as each new series was introduced." (p. 502)

These assertions were based on observations of his unmusical listeners, verbally assigning frequency rates to pure tones, with an average error of approximately 3 semitones.

Komatsu (1940)

The first account of experimentation on training absolute pitch by a Japanese author seems to have been that of Aiko Komatsu. In 1940 she published her results in a Japanese journal devoted to the study of educational psychology. Komatsu reported that it was widely believed in Japan that absolute pitch could not be acquired after birth, and therefore music education experiments had, for the previous 60 years, dealt only with the acquisition of relative pitch. Komatsu stated her belief that there are three distinct levels of absolute pitch, although they need not necessarily be developed according to that order. The lowest level of absolute pitch concerns the ability to identify single tones. Komatsu felt that this level has
the least musical value of the three. The middle level concerns absolute pitch identification of chords. Komatsu wrote that some possessors of this skill could name consonant and dissonant chords out of context, but they could not name the pitches of single tones. In any event, this achievement is not yet musically sufficient. She deemed the highest level of absolute pitch to be the ability to identify the keys of pieces of music, in addition to possessing the skills of the two lower levels. Komatsu’s training experiment attempted to determine which, if any, of these three levels of absolute pitch could be taught to a child who had never received previous musical training.

Using a young girl less than 5 years old as her subject, Komatsu began her absolute pitch experiment in September, 1938. The child’s father was a musician who played the flute; the child’s mother was a school teacher who played the piano a little bit. Neither parent possessed absolute pitch. The little girl was the youngest among three sisters. Her two older sisters had each taken piano lessons, but the test subject herself had not received musical training formally or informally. Her personality was described as strong-minded and typically childlike. The training regimen devised by Komatsu encompassed only the middle level of her absolute pitch hierarchy. Twice each week, on Mondays and Thursdays, she played triads on the piano for the subject at a mezzo forte volume level, with a marcato touch. After each presentation she read, in German, the names of the tones, from lowest to highest for the subject to repeat. The length of each session was not specified. By March, 1939, the training had progressed to include 25 triads of different types, and these triads were drilled during April and May. From June to the beginning of November, seventh chords were added, raising the total number of different chords to 46. This completed the chord training, although
Komatsu stated the intent to introduce dyads, and single tones, at a later time. The black keys of the piano were to be taught following mastery of the white, but the use of movable do solmization was prohibited throughout the training, and the cooperation of the family was requested in this matter. Review of the training was to be completed at home, and this duty was assigned to the subject's mother.

After about 15 months of training, when the subject was 6 years and 1 month old, she could name the pitch of the root and the quality of a chord immediately upon hearing it, and could also indicate her answer in music notation. While her response times ranged from 0.5 seconds to 1.0 second, incorrect responses required more time than correct ones. Komatsu reported that during the entire training period the subject's percentage of correct responses might vary from a low of 64% to a high of 96%. She did not report the details of the testing procedure, presumably because they were the same as those of the training procedure. The subject's ability to identify chords was not influenced by timbre, for she responded as accurately to chords on a record as she did to those presented on piano. Komatsu reported that every time the subject had stayed up too late the night before, or played too much the day before, her accuracy suffered. Interestingly, the subject always responded in a low-pitched voice to low-register piano chords, and in a high-pitched voice to ones of the upper registers. Komatsu also noted that the written response task was still difficult enough for the subject to cause frequent notational errors.

After the chord training was completed, as an informal study, Komatsu presented five or six isolated tones (among them C4 and A4) to the subject. Noting that her findings were only preliminary, she reported that the subject
was able to identify these tones correctly, even though she had never received any training of this type. Komatsu wrote that, to date, this experiment was just part of a program planned to continue for many years, but that she now assumed that absolute pitch could indeed be developed even in those who lack musical training. No reports of further experimentation by Komatsu have come to light.

While the acquisition and translation of this report of Komatsu’s experiment represents a novel and fascinating addition to the English language literature on absolute pitch, the absence of vital experimental details casts doubt on its merit and usefulness. Specifically, Komatsu’s report did not identify the different types of triads used in the training. There is no way of knowing which registers they were presented in, whether the presentations were arpeggiated or blocked, or whether any were presented in inversion. Likewise, no details are provided concerning the testing procedure(s) used. It can only be hoped that Komatsu did not determine the percentage of correct responses from raw data derived solely from the training sessions, where the use of relative pitch would be possible. Additionally, there is no reported measure of the subject’s ability prior to the commencement of training, or indication of the lengths of the training sessions. When one considers that Komatsu used the smallest possible number of subjects with no established control group, meaningful interpretations of her data become difficult, if not impossible.

**Hindemith (1949)**

To the previous attempts by composers Jadassohn and Maryon to teach absolute pitch, we must add yet another by the celebrated Paul Hindemith
(1895-1963). Hindemith’s *Elementary Training for Musicians* (1949) represents a comprehensive attempt to combine the rudiments of music theory (rhythm, meter, intervals, scales, notation) and aural training (dictation, sight-singing, sight-reading, clef-reading) into a single, concise manual. The book was designed to train students aspiring to become professional musicians, and sought to raise them to a level suitable for entry into a beginning-level college harmony class. It does so by continuously integrating the presentation of relevant theoretical information and student activities. The manual is laid out in two parts: Part One contains 11 chapters devoted to musical rudiments, and Part Two contains related dictation exercises. The only new topic introduced in the dictation portion is Hindemith’s coverage of absolute pitch.

Hindemith’s dictation exercise 21 is designed to accompany Chapter Five, which introduces dotted notes and dotted rests, the small octave (the conventional Helmholtz designation for the third octave in USA Std. system), the major scale, sharps and flats, diatonic and chromatic semitones, and key signatures. Exercise 21 (1) is a simple drill wherein students practice making correct octave designations by the use of the Helmholtz system. Example 21 (2) introduces Hindemith’s discussion of absolute pitch. It simply reads “TEACHER: Play tones within the range e—c” [E₃—C₆] on the piano. PUPIL: Name them.” (Hindemith, 1949, p. 206) Hindemith stated that while this task would pose no problem for those with absolute pitch, every normally gifted musician can be trained to acquire this skill. Hindemith (1949) wrote:

> The ear of every normally gifted musician can be trained to become what is supposed to be “absolute” in its judgment, the more so since there is no such absoluteness, based on any physical attribute of the ear.
What we call “absolute pitch” is merely a highly developed ability to compare quickly an audible impression with acoustic archetypes stored in our memory. These archetypes are always taken from earlier musical experience and are closely related to all kinds of external qualities of tones and of tone-production. (p. 206)

Hindemith cited his own experience as evidence that absolute pitch could be learned, and mentioned the lack of a universal fixed-standard for pitch as evidence to support a learning theory. Hindemith believed that the absolute pitch of the violinist, for example, likely depended upon the timbre and register of his instrument. Likewise, the absolute pitch of a singer might be based upon the muscular tension of his vocal cords.

Hindemith mandated that all his dictation exercises begin with the sounding of an a on the tuning fork. Beginners can practice exercise 21 (2) without great leaps between tones, with frequent reference to the tuning fork, and, if need be, with leaps being approached with step-wise motion from the a. However, at this stage of training, he hoped that students can remember the tuning-fork pitch whenever they heard another a. Thus, he proposed that each exercise should now start by singing a before sounding the tuning fork, instead of sounding the tuning fork first as had been done previously.

Hindemith (1949) wrote:

This experiment may at first fail frequently enough, but after eighty or a hundred attempts a fairly firm and reproducible impression of a must be established. If not, the question may be raised whether there is any musical gift at all in the mind that cannot learn to remember and compare pitches. (p. 207)

Pollack (1952)

Irwin Pollack, of the Communications Division of the Human Resources Research Laboratories of the U.S.A.F., was the first investigator to apply information theory to subjects’ absolute identifications of the pitch of
pure tones. Pollack (1952) investigated the discrepancy between listeners' abilities to make fine discriminations among pitches (just-noticeable differences) and their relative inability to identify them. Instead of arbitrarily deciding upon an acceptable level of subject error, Pollack calculated the bits of information that are transferred by subjects in the process of identifying a series of individually presented tones, in order to arrive at the number of perfectly identified tones that would constitute an equivalent level of performance.\footnote{The unit of measure used in information theory is the \textit{bit}. The number of bits of information transferred to the subjects in this study is equal to the logarithm (base 2) of the equivalent number of tones perfectly identified. The correct identification of four different tones would yield two bits of information. The correct identification of eight different tones would yield three bits of information, and so on.}

Pollack manipulated only one variable in each series of tones, their frequencies. Using a loudness level of approximately 85 dB, he presented sine tones randomly, but with an equal number of presentations of each tone per session. Each session involved a different series of tones, with each successive series adding more. Regardless of the total number of tones in each series, frequencies followed equidistant logarithmic divisions from 100 Hz to 8000 Hz. The duration of each stimulus tone was approximately 2.5 seconds, and there was about a 25-second pause between presentations to allow subjects to identify the tone and record its assigned number on an answer sheet. Pollack (1956) identified this method as "the psycho-physical method of absolute judgment." (p. 746) He went on to state that "after the subject entered his response, the correct identification was supplied before the next tone was presented." If he followed this procedure in test sessions as
well as in practice sessions, it is unlikely that his subjects made an absolute judgment in any but the first response of each session.

Pollack's first experiment began by testing groups of up to 38 subjects responding to series of 2, 3, 4, 5, 6, 7, 8, and 10 stimulus tones reproduced by a loudspeaker in an untreated room. The information transferred to subjects increased as the number of tones increased until a plateau was reached at a level of 2 bits per stimulus presentation. This was equal to the perfect identification of only four different tones. A second experiment attempted to increase this number by testing the subjects individually and by presenting the stimuli via headphones. Pollack wrote that in this experiment "testing was performed with a group of more experienced subjects ... and their results considered only after several days of practice testing." (Pollack, 1952, p. 746) (While this is an example of short-term training, it was deemed significant enough to be included in this discussion of training attempts.) These changes brought about slightly improved results, and the judgments of these 6 subjects on series of 2, 3, 4, 5, 6, 7, 8, 10, and 14 stimulus tones yielded an average approaching an asymptote of approximately 2.2 (pooled) or 2.3 (unpooled) bits of information transferred. This number approaches the number required for the perfect identification of five tones (2.32 bits). The lack of improvement that he found in subjects' responses when he increased the number of alternative tones beyond five confirms Hake & Garner's (1951) conclusion that an increasing in the number of alternatives beyond the

---

15 Pollack used average (unselected) listeners, not musicians, as subjects. He reported that none possessed absolute pitch.
16 Indeed, cognitive psychologist John R. Anderson broadly asserted that "it requires at least 100 hours of learning and practice to acquire any significant cognitive skill to a reasonable degree of proficiency." (Anderson, 1982, p. 369)
maximum that can be perfectly identified will not increase the amount of information transmitted.

Noting that uncertainty among subjects increased in the middle of the frequency range, centered on approximately 1000 Hz, Pollack partitioned the frequency space into high, middle, and low registers. Using eight tones, he systematically omitted registers and explored relatively open and close spacings in the hope of increasing the amount of information transferred by the series. He found, however, that no scheme of partitioning the frequency space yielded an informational transfer greater than 2.1 bits per stimulus.

In Pollack's final attempt to increase the information transferred to subjects, he varied the upper limit of the frequency range from 8000 Hz to 500 Hz while maintaining the lower limit at 100 Hz because he thought the frequency range was perhaps too large for efficient encoding. Eight-tone series were presented, and, in each case, were still spaced at equal logarithmic intervals with 500, 1000, 2000, 4000 and 8000 Hz upper limits. Surprisingly, while this span represents nearly a twenty-fold reduction in frequency range, the information gained varied by only about 0.2 bits, or roughly 10%. In fact, the information gained by subjects fell rather steadily as the range was reduced—from a high of approximately 2.0 bits for the 7600 Hz range to a low of approximately 1.75 bits for the 400 Hz range.

Pollack came to the conclusion that, bound by his stimulus restrictions and the conditions used in informational measures, response accuracy in the absolute judgment of tones in a series is relatively invariant, regardless of the numbers of tones, their frequency range, or their distribution. Even though his best subject could occasionally identify seven tones correctly (almost 2.8
bits), the average informational transfer obtained by Pollack's subjects was 2.3 bits, or perfect identification of a maximum of five tones.

In 1953 Pollack conducted a trio of follow-up experiments. Because they neither represent an attempt at training absolute pitch nor include subjects already possessing absolute pitch, no discussion of them is merited here.

**Hartman (1954)**

E. B. Hartman of the U.S.N. Submarine Base in New London, Connecticut, was familiar with Pollack's recent experiments, and with the attempts of Meyer, Köhler, Gough, and Mull to train absolute pitch. Hartman (1954) was aware that the latter experimenters' failure to control important variables had rendered their results suspect. He cited the number of tones in the series, the interstimulus time-interval, and the pitch distance between tones as examples of these variables. To test rigorously the limits of the absolute judgment of pitch, he held constant the number of tones in the series, while varying the distance between the tones. He further tested the effect of practice on the acquisition of this new skill, and its atrophy following a period of relative disuse.

Hartman divided a subject population of 12 according to their performance on an absolute pitch pretest, requiring the identifications of 36 sinusoidal semitones between C3 and B5. While individual pretest scores were not reported, we know that four equivalent groups of 3 subjects were extracted in such a way that each group had an average error of approximately 4 semitones. Each group trained using a single series of nine pure tones that
were separated by 50, 100, 200, and 300 mels, respectively. The nine-tone series were apparently centered on 1700 Hz, and obviously had vastly different ranges. Tones were generated on a General Radio Company Interpolation Oscillator, recorded on a phonograph record, and replayed to subjects listening through headphones in a quiet room. Each tone was presented at a loudness level of approximately 50 phons as determined by a graph of equal-loudness contours and the tones were arranged in series wherein each tone followed every other tone an equal number of times. Subjects were presented a 6-second burst of white noise in an attempt to erase previous tones from their memory. Then, after a 1-second silent interval, they were presented the 2-second pure tone. Subjects were instructed to answer quickly, and to make no attempt to remember the previous tone while awaiting the next. This sequence was repeated every 30 seconds. Subjects identified the tones by writing one of nine assigned numbers visible to them on a chart. The numbers ran from low to high corresponding to frequency. Subjects trained twice a week for 8 weeks, and each session provided eight repetitions of the nine-tone series. Hartman gave corrections after every response for the first 7 weeks of training; none was provided during week 8.

Though progress was made by all four groups, no group was able to master perfectly its nine-tone series in the 2-month period. During the first

---

17*American National Standard Psychoacoustical Terminology*, a publication of the American National Standards Institute, defines the mel as “a unit of pitch. One thousand mels is the pitch of a 1000 hertz pure tone for which the loudness level is 40 phons. The pitch of a sound that is judged to be \( n \) times that of a 1000-mel tone is \( n \) thousand mels; thus, the pitch of a sound subjectively judged to be \( n \) times that of a 1-mel tone is \( n \) mels.” (1973, p. 27)

18Except Session 10.
week of training the mel-spacing factor was apparently unimportant; none of the groups could obtain more information from their nine-tone series than they could from the perfect identification of two or three tones. The group averages ranged from 1.0 to 1.3 bits of information transmitted. However, as training progressed the groups began to diverge. By the end of training, the 50-mel group had improved from approximately 1.05 bits to approximately 1.25 bits. Unfortunately, this was still in the range of the amount of information equivalent to the perfect identification of only two or three tones. Similarly, the 100-mel group steadily improved (though their initial progress was greater than that of the 50-mel group) and the level of information transferred rose from an initial level of approximately 1.0 bits to almost 1.4 bits. Again, this was still in the range of the amount of information equivalent to the perfect identification of only two or three tones. In both cases the effect of training was hardly significant. The results of the two remaining groups were more encouraging. The 200-mel group improved from an initial measure of approximately 1.15 bits to slightly more than 1.6 bits. This yields an informational measure equivalent to the perfect identification of three or four tones. The 300-mel group experienced the steepest learning curve of all of the four groups, and rose from an initial level of 1.3 bits, to a final level greater than 2.2 bits (almost a full bit more than the 50-mel group). The amount of information contained in this measurement is equivalent to the perfect identification of four of five tones, a performance comparable to Pollack’s findings.

In summary, the amount of information transferred varied directly with the size of the separation of the tones in mels. For the subjects experiencing smaller amounts of separation between the tones (50 and 100
mels), the learning curves were generally flat and the slight progress was made fairly slowly. For the subjects experiencing larger amounts of separation between the tones (200 and 300 mels), the learning curves were generally steep and significant progress was made quickly, by comparison. Hartman also found that after 2 months of rest from the absolute judgment of pitches, all four groups of subjects showed a decrease in abilities. The amount of the loss was generally inversely proportional to the size of the separation of the tones in mels. The 300-mel group's performance was affected very little; the 200-mel groups' performance fell about 0.2 bits, while the 100-mel group's fell slightly more. The single exception was the performance of the 50-mel group. It fell by only about 0.1 bit, but that is not surprising given the modest success that these subjects had in acquiring the skill at all. Hartman concluded that in the absolute judgment of tones, pitch spacing was positively related to the number of tones that could be learned, the rapidity of learning, and the long-term retention.

Deutsch (1956)

Yet another attempt to train absolute pitch by a prominent musician is that of trumpeter and composer Maury Deutsch (1954). Deutsch began playing trumpet at the relatively late age of 15, but in 1 1/2 years he had progressed so remarkably that he was awarded the prestigious New York Philharmonic Symphony scholarship which enabled him to study with renowned trumpet pedagogue Max Schlossberg. Deutsch earned a doctorate in music from the Musical Arts Conservatory, Amarillo, Texas and is the author of numerous books on arranging, composition, ear-training and
conducting. Deutsch firmly believed that absolute pitch could be learned and that he could teach it.

While Deutsch declared that "it is not difficult to differentiate pitches by comparing the relative changes of tension in the vocal cords" (1954, p.16), he thought that true absolute pitch was a mental process that is not dependent on other faculties. In his article he provides preliminary exercises for the student to practice prior to undertaking his method. The first true objective is to have the student "visualize" a single tone so firmly that it can be recalled at will. The choice of the specific pitch is left to the individual, but it should be one that is easy to sing and "visualize." Students are instructed to play the tone, listen to it while attempting to "visualize" it, repeat the process after 30 seconds, and check to see if their memory has drifted. This exercise is to be practiced a minimum of five times daily. Next, after the successful "visualization" of the tone, the student must concentrate on its letter name in the hope that, after enough repetitions, the thought of just the letter name will evoke the tone's "visualization." Deutsch likened this first step to Pavlov's famous conditioning studies.

After an internal standard has been established, additional tones are added. The succeeding tones are separated by relatively large dissonant intervals in order to impede the use of relative pitch. Specifically, Deutsch recommended the second tone lying an augmented fourth from the standard, the third tone either a major 7th or a minor 9th from the second, with succeeding tones following the same arrangement as shown in Figure 1.
Figure 1. An example of the order of memorization for a series of tones, arbitrarily beginning on $E_b$, recommended by Deutsch (1954).

The distinction between the ascending major seventh and the descending minor ninth is apparently unimportant, because they both are considered equivalent using modulo-12 arithmetic. Of course, the direction of the tritone is trivial, because it remains invariant under inversion. The appropriate placement of the starting tone in the student's vocal range is of primary importance. New tones are to be added only after achieving mastery of old ones, and students' current series is to be practiced daily by reviewing the tones with a minimum time lapse of 10 minutes separating each one. This is done to minimize the effects of learning the serial positions among the tones. After the successful memorization of these 12 tones, the next step is to "visualize" these tones in other octaves.

For students not wishing to devote the time needed to acquire all 12 pitches, Deutsch recommended a simulated form of absolute pitch whereby students could simply memorize one, two, or three pitches, and determine the others by means of relative pitch. The selected tones should center in the middle of the student’s vocal range. Tones selected by Deutsch himself consist of two successive ascending minor sixths and form a broken augmented triad. This pattern represents an equal division of the octave whereby any potential tone to be named could lie no further away than a major second from a familiar pitch-class.
Figure 2. Deutsch’s three-tone groups to be memorized in a simulated form of absolute pitch by (a) sopranos, (b) altos, (c) tenors, and (d) basses.

Additionally, Deutsch believed that his “visualization” technique could be successfully applied to the memorization of absolute pitches used in occidental scales, or equal tempered scales of 5- or 24-to-the octave. Although Deutsch provided no experimental or anecdotal evidence regarding the success of his method, he must surely have had faith in these techniques, inasmuch as he published almost verbatim, under a slightly different title, the same article some 32 years later (Deutsch, 1986).

Lundin & Allen (1962)

Hamilton College psychologists Robert Lundin and Joseph Allen joined forces in an attempt to establish criteria sufficiently stringent to distinguish pitch discrimination from absolute pitch, and tested the latter’s development under training. They (1962) wrote:

When a subject is presented with a random series of musical tones within certain limits of the musical scale, if he can identify them correctly without reference to other objectively presented and already named pitches, we will say he is exhibiting the behavior commonly designated as perfect pitch. Accordingly, if we can contrive an
experimental situation which meets this criterion and if we can train one or more individuals so they can name pitches correctly when these are presented randomly and without prior reference tones, we believe we have trained them in perfect pitch. (pp. 139-140)

Lundin & Allen’s experiment began with a population of 5 subjects, all of whom were professors or undergraduates at Hamilton College. Two of the subjects claimed to possess absolute pitch already, 2 had “substantial musical training,” and 1 had no musical training—he could neither read music nor play an instrument. The training used techniques similar to those employed by teaching machines in the programmed learning of verbal materials. The apparatus included a reel-to-reel tape recorder to play tapes of single tones produced on a recently tuned Baldwin grand piano (A4 = 440 Hz). Eight tapes were made, each containing a unique random ordering (without replacement) of the 24 chromatic pitches ranging from C4 to B5. Subjects were not informed, however, that a pitch would be heard only once in any given series. Subjects listened to the tapes through headphones while seated in front of a response panel in a soundproof booth. The response panel had 24 buttons, laid out in two vertical columns of 12, and each button activated a light on a large musical staff placed approximately 10 feet from the subject. The staff was visible to observers through a glass window in the booth. Only sharp symbols were used to designate chromatic pitches, and provisions were made to distinguish clearly the two pitches that occupy the same line or space.

19 In contrast to this description of the subjects (Lundin & Allen, 1962, p. 140), Lundin himself later wrote that “In our original training procedure we used five male college students as subjects...” (Lundin, 1963, p. 49)
20 The specific deck used had a frequency response of 40 to 15,000 Hz at 7.5 inches per second, and a signal to noise ratio 48 dB down from 3% at the third harmonic (a once-common measure).
In an introductory session, each subject became familiar with the apparatus; tones were played and the subject was allowed to explore the relationship between the response panel buttons and the lights on the metal staff representing musical notation. The subjects’ second session was a pretest of pitch naming ability, and they responded by pressing the buttons on the control panel. After the presentation of a tone, the subject would request the next tone by means of a switch on the response panel. Although the presentation rate in the pretest was apparently determined by the will of the subject, the subject received no evaluation of his responses.

The training sessions were conducted in much the same manner except that feedback was given. Immediate feedback was provided when the subject flipped an answer switch on the response panel. A correct response appeared brightly; if the response was incorrect, both it and the correct answer remained lit until the subject pressed the correct button. This procedure indicated not only the correctness of the response, but also the number and direction of an error in semitones. Such information would obviously be the least useful to the musically naive subject. As for the presentation rate of the training sessions, Lundin & Allen stated that “about 20 seconds elapsed between each tonal presentation, making it difficult, if not impossible, for the subject to use the previous stimulus as a reference tone.” (Lundin & Allen, 1962, p. 142) This assurance is highly suspect, however, inasmuch as studies investigating the persistence of tones (i.e., Rakowski, 1972) show that retention lasts for much longer periods of time.

Subjects endeavored to train daily, though no indication is given as to the length of each individual session. Tapes were randomly selected, except that the same tape was never permitted to repeat itself and subsequent
presentations were separated by a time interval to prohibit its memorization. Three of the 5 subjects completed the full regimen of 36 training sessions, 1 completed only 17, and the other, the musically naive subject, dropped out after completing only 11 training sessions. Thus, only the 4 remaining subjects were administered a posttest of 24 random tones presented as in the pretest, without feedback.

One of the conclusions reached by Lundin and Allen was that the 2 subjects who claimed to possess absolute pitch prior to training were mistaken; their percentages of correct responses on the pretest were only 54% and 42%. Another conclusion that bears more directly on the purpose of the study was that all 4 subjects who completed the posttest showed significant improvement in their pitch naming ability, compared to the levels shown on the pretest. The 2 subjects who claimed to possess absolute pitch attended all 36 training sessions and improved their correct response rates from 54% to 100%, and 42% to 96%, respectively. Of the 2 subjects who had “substantial musical training,” one improved his correct response rate from 17% to 42% after 36 training sessions, and the other from 4% to 17% after 17 training sessions. The musically naive subject had an 8% correct response rate on the pretest; his progress during 11 training sessions was not adequately reported. In summary, the average percentage of correct responses for 4 of the 5 original subjects rose from 29.25% on the pretest to 63.75% on the posttest. Table 4 provides a partial summary of Lundin & Allen’s data.
Table 4. Summary of individual subjects’ data for Lundin & Allen (1962) showing a) the percentage of correct responses on the 24-tone pretest, b) the percentage of correct responses on the 24-tones posttest, c) the percentage correct responses during the first four training sessions, d) the percentage correct responses during the last four training sessions, e) the percentage improvement from the first four training sessions to the last four training sessions, f) the average error in semitones for the first four sessions, g) the average error in semitones for the last four sessions, and h) the total number of training sessions each subject attended.

<table>
<thead>
<tr>
<th>Subject</th>
<th>a</th>
<th>b</th>
<th>c</th>
<th>d</th>
<th>e</th>
<th>f</th>
<th>g</th>
<th>h</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Claimed AP</td>
<td>54</td>
<td>100</td>
<td>72</td>
<td>95</td>
<td>23</td>
<td>1.5</td>
<td>1.0</td>
<td>36</td>
</tr>
<tr>
<td>2. Claimed AP</td>
<td>42</td>
<td>96</td>
<td>56</td>
<td>91</td>
<td>35</td>
<td>1.1</td>
<td>1.0</td>
<td>36</td>
</tr>
<tr>
<td>3. Musically trained</td>
<td>17</td>
<td>42</td>
<td>27</td>
<td>60</td>
<td>33</td>
<td>2.1</td>
<td>1.6</td>
<td>36</td>
</tr>
<tr>
<td>4. Musically trained</td>
<td>04</td>
<td>17</td>
<td>27</td>
<td>46</td>
<td>19</td>
<td>2.9</td>
<td>2.0</td>
<td>17</td>
</tr>
<tr>
<td>5. Musically naive</td>
<td>08</td>
<td>—</td>
<td>24</td>
<td>35</td>
<td>11</td>
<td>2.4</td>
<td>2.2</td>
<td>11</td>
</tr>
</tbody>
</table>

Lundin and Allen concluded that one subject had met their perfect identification criterion for absolute pitch, and a second subject was extremely close, missing it by only one judgment. All subjects made significant gains in their pitch naming accuracy, but the increase was smaller for the 2 subjects who failed to complete the series of training sessions. Additionally, a decrease in the average size of the errors was reported for all subjects following training; this, of course, was largely a factor of the individual subject’s room for improvement. Error analyses were completed for only 1 of the subjects, Subject 3 (refer to Table 4). They showed that this subject identified the tones of the C diatonic collection more accurately than chromatic tones, and the tones ranging from C4 to B4 were more accurately identified than those from ranging C5 to B5.

---

21 An accuracy level of 96% correct judgments is the equivalent of 23 correct out of 24 posttest responses.
Lundin (1963)

In 1963, Lundin, who had recently become a member of the Editorial Board of the *Music Educators Journal*, published an interim report of a second experiment that he was conducting with Allen (Lundin, 1963). The experimenters were attempting to determine if the training of absolute pitch would be facilitated by applying the techniques of programmed learning. First they presented the most easily identified tones, as determined in their first experiment, and gradually added the more difficult tones.

Five new subjects were selected, all of whom had received “some musical training prior to experimentation.” (Lundin, 1963, p. 51)\(^{22}\) They trained in a manner similar to that of the first experiment except that they heard 14 tapes of graduated difficulty, instead of the 8 equally difficult tapes previously used. The first tapes contained only tones from the C diatonic collection. Chromatic tones were gradually introduced on succeeding tapes according to the standard order of the sharps.\(^{23}\) In the intermediate tapes, individual tones might appear more than once per tape; in the final tapes, subjects were presented all 24 chromatic tones ranging from C₄ to B₅, randomly ordered without replacement, as in the first experiment. A new aspect of this experiment was that subjects were permitted to train at their own pace.

Lundin reported that although individual differences in the rate at which learning took place were apparent, all subjects made “remarkable

---

\(^{22}\)Other available details reveal that a pianist, a clarinetist, and a singer were among the five subjects.

\(^{23}\)F♯, C♯, G♯, D♯ and A♯ were all that were required to complete a chromatic collection.
improvements.” Regarding the improvement in the number of correct judgments, he wrote:

Introducing a “test” in the middle of training (that is a trial in which no information or feedback is given) we found, for example, that after 10 trials, one subject had made 700% improvement, another 600%, another 200%, and so on. It is evident, of course, that in the beginning these subjects who showed such marked improvements started out with very few correct responses.24 (p. 51)

Lundin also reported the average size of errors greatly decreased, citing in particular the performance of one subject who, in 120 responses, made only 4 errors greater than a semitone. It is unfortunate that complete pretest and posttest data, either individual or group, were not provided in his report, but as stated earlier, at the time of the report the experiment was still ongoing. A final report of the study has not been located.

Terman (1965)

In 1965, Michael Terman of Brown University published the details of his training experiment in absolute pitch. Terman’s work had been conducted at Columbia University, utilizing the laboratory facilities of Cyril Harris and Vladimir Ussachevsky. Terman had studied Lundin and Allen’s (1962) training experiment, and sought to improve upon their methodology. He objected to their use of a 20-second interstimulus silent interval, noting that “stimulus tones may be recalled and reproduced with high accuracy after 20 sec. of intertrial silence, and thus used as relative cues, weakening the ‘absolute’ construct.” (p. 243)

24For a subject to have made a 700% improvement, the original number of correct responses could not have been greater than 3 out of 24. If fact, Lundin did report that two of the five subjects made only one out of 24 correct responses.
Terman used as subjects 11 adults having “various musical backgrounds.” None claimed absolute pitch. These subjects trained using seven tapes on TMI-Grolier Min/Max teaching machines that were modified to provide correction after each response. Each tape contained 25 trials of randomly selected sinusoidal stimulus tones ranging from C3 (130.8 Hz) to C5 (523.3 Hz) inclusive (25 equal-tempered semitones). Terman interspersed brief periods of silence, “random” fire-siren-like sine tone sweeps (whose frequencies extended above and below the test range), and subject vocalization in an attempt to minimize the effect of intertrial cues.

Each trial proceeded as follows. First, loud sine tone sweeps faded in and lasted for 10 seconds. Then, there were 15 seconds of soft sine tone sweeps presented in the background, while the subject was directed to read aloud. Next, there was another period of loud sine tone sweeps, but this time for only 5 seconds. This was followed by 10 seconds of silence. Finally, the sinusoidal stimulus tone was presented for 3 seconds at approximately 92 dB SPL. During a 20-second silent period before the next sequence began, subjects responded in writing, identifying both the octave and pitch class of the stimulus tone and the teaching machine, noting the response, and providing correction, if necessary. Subjects completed two training tapes per week in an hour-long training session. Thus, though it was not explicitly stated, it appears that because of repetitions the training required 14 weeks to complete.25

---

25This estimate was made using Terman’s plots of some representative individual learning curves. The abscissa consists of seven data points labeled “Training tapes (Blocks of 4)” (p. 244) [7 x 4 = 28; 28 ÷ 2 tapes per week = 14 weeks] This corroborates a brief reference that “individual tapes were not presented consecutively, but were repeated three to five times in the course of the experiment.” (p. 243) [7 tapes x 3 presentations each = 21; 21 + 2 tapes per
Terman found that 10 of the 11 subjects showed improvement after training. On the final test, his highest-scoring subject achieved a score of 82% correct responses. Moreover, approximately 95% of his responses were within a semitone of the correct answer. Before training, this subject's score was only about 50% correct. Terman also found significant differences in the progress of other individuals. For example, 2 subjects, who each began training with mean percentage scores between 10-15% correct, achieved terminal proficiencies of 38% and 75%. Terman reported not finding any apparent relationship between subject improvement and musical experience, but gave no details.

Terman acknowledged some shortcomings in his own experimental methodology. Upon examination of the group data for the last eight training tapes he stated:

Pitch naming accuracy was highest in the extremes of the stimulus continuum, and a roughly U-shaped function connects the extremes, suggesting that relative accuracy for naming the various stimuli was partially a function of a procedural artifact—the limited range of test frequencies. (p. 243)

He also acknowledged that although 10 of the 11 subjects found the distinctive procedures effective enough to prevent their remembering stimulus tones, one did not. Unfortunately, the subject who reported remembering stimulus tones through the distinctive procedure and who claimed they provided relative cues was the overall high-scorer.

---

week = 10.5 weeks minimum, at the other extreme, 7 tapes x 5 presentations each = 35; 35 + 2 tapes per week = 17.5 weeks maximum]
Fisher (1965)

Fred Fisher, a graduate of the Eastman School of Music and faculty member of Oklahoma State University, wrote in 1965 from a personal perspective about the effect of training on the development of absolute pitch. Fisher, the son of conductor-composer Irwin Fischer, recounted both indirect and direct aspects of early parental training that resulted in his acquisition of absolute pitch.

Fisher (1965) recalled that his mother taught elementary piano lessons in their home as he and his brother were growing up:

In retrospect, I can still recall the incessant sound of middle c—the note on which all beginners' pieces seem to start. The first piece went something like this: c c c c /c — c — /c c c c /c — . Then this phrase was repeated with the other hand. Sometimes there were c's too many, sometimes too few; sometimes the rhythm was wrong, and sometimes it stopped altogether. But I'm sure that in the space of a very few years the sound of middle c was as familiar to me as my own name. (p. 18)

Fisher also recalled the music-making activities of both his parents and their more advanced students, and tells how he and his brother discovered which piano keys would reproduce the piano sounds coming from another room.

More directly, Fisher told of how both parents worked to instill in him a sense of absolute pitch. As early as the time when he and his brother were learning to walk, his parents attached small gold stars on the fronts of all the c's on the family piano. At first, this attraction led the children merely to sound these tones. Later, they were asked to name them, and then to learn the relationships between c's and the other tones. As additional exercises, their father would sound groups of tones and ask the children to name them. Tones were sounded simultaneously or in succession, and these little drills were conducted frequently. While the children benefited from frequent
training and enjoyed great parental support, Fisher’s account (1965) also states that “a private piano teacher of my acquaintance was able to develop pitch recognition in eight out of fourteen children of elementary school age without significant help from the parents.” (p. 19)

For the development of absolute pitch in older students, Fisher suggested that they begin with just one pitch, as he had done as a child. The first drill he recommended was that the student gradually try to lengthen the recall time for a tone. In early tries, the student might check himself after an hour; later, the interval might be lengthened to a full day or overnight. Fisher’s second drill called for a partner to sound a series of four or five unrelated tones, the only requirement being that one of them must always be middle C. In early attempts the tones should be widely spaced tones, and as the student progresses, the tones can be closer together, but not so close that the student becomes confused. Fisher specified that this exercise be practiced at hourly intervals, yet never when the student is tired. He stated that while the results of this training may not be immediately discernible, the student’s progress should be measurable within a few weeks.

Fisher stated further that while absolute pitch can be learned, it can also atrophy through disuse, and he offered a first-hand example. He and his brother both learned absolute pitch as children, but the long term result proved different. Fred pursued a musical career and retained his ability while his brother became an engineering student, devoted little time to music, and lost his.
Vianello (1966)

Michael Vianello conducted a thesis study, published in 1966, at Texas Christian University under Dr. Selby Evans. Vianello wished to test, like Pollack (1952) and Hartman (1954) before him, the apparent differences between pitch discrimination and pitch identification. Citing Miller's (1956) theoretical limit regarding the number of pitches that can be identified in absolute discrimination tasks (7 plus or minus 2, or about 2.8 bits of information transmitted), Vianello measured the effect of training on subjects who were trained at first without feedback, and later with feedback.

Vianello began with 30 college students as subjects, some of who received credit in introductory psychology courses for their participation. Only 22 subjects completed the experiment. No description of their previous musical training or current musical abilities was given in his report, or in a later summary of the experiment (Vianello & Evans, 1968). Stimuli consisted of 18 sinusoidal tones, each with a duration of 3 seconds; the interstimulus time interval was 7 seconds. While the later report stated that the tones were "arranged in three scales with frequencies corresponding to the musical notes A, B, C, E, F, and G" (Vianello & Evans, 1968, p. 576), inspection of the original report (Vianello, 1966) revealed that five chromata, not six, were used. The frequencies of the tones used were 220, 240, 272, 348, and 440 Hz, along with their multiples at the double and quadruple octaves. Although these frequencies may "represent" the pitches A₃, B₃, C₄, F₄, G₄, and A₄, not all are the frequencies used in equal temperament.²⁶ Subjects responded by identifying the pitch of each tone, using a system of coded alpha-numerical

²⁶Commonly accepted frequencies would be 220.0, 246.9, 261.6, 349.2, and 440.0 Hz.
symbols. Under Vianello's encoding system, the frequency of 220 Hz was indicated by the symbol A1, 880 Hz was indicated by A2, and 3520 indicated A3.\textsuperscript{27} Table 5 shows the signal and label array used by Vianello (1966).

**Table 5.** The response labels associated with frequencies, in Hertz, of stimulus tones utilized by Vianello (1966).

<table>
<thead>
<tr>
<th>&quot;Octave&quot;</th>
<th>&quot;A&quot;</th>
<th>&quot;B&quot;</th>
<th>&quot;C&quot;</th>
<th>&quot;F&quot;</th>
<th>&quot;G&quot;</th>
<th>&quot;H&quot;</th>
</tr>
</thead>
<tbody>
<tr>
<td>&quot;1&quot;</td>
<td>220</td>
<td>240</td>
<td>272</td>
<td>348</td>
<td>384</td>
<td>440</td>
</tr>
<tr>
<td>&quot;2&quot;</td>
<td>880</td>
<td>960</td>
<td>1088</td>
<td>1392</td>
<td>1536</td>
<td>1760</td>
</tr>
<tr>
<td>&quot;3&quot;</td>
<td>3520</td>
<td>3840</td>
<td>4352</td>
<td>5568</td>
<td>6144</td>
<td>7040</td>
</tr>
</tbody>
</table>

Subjects trained in three sessions per week for 2 weeks. In sessions One through Four, participants were asked to identify a total of 756 test tones plus 72 warm-up tones. No feedback was provided during these sessions. In Session Five, subjects were asked to identify 216 tones plus 18 warm-up tones. Following each response, the experimenter now provided the correct pitch label. The format for Session Six, the final session, was the same as for Session Five, except that no feedback was provided. The results of Session Six were then compared to those of Session Four to determine the effect of training with feedback (Session Five).

Vianello reported that average subject improvement increased from 2.17 bits of information transmitted in Session One to 2.38 bits in Session Six. The author surmised that while performance varied across subjects, absolute pitch identification ability improved as a function of training without feedback, and that subsequent training with feedback resulted in further improvement. He also reported that performance differs widely across

\textsuperscript{27}The authors offer no explanation for their decision not to use U.S.A. Standard octave designations, or for their failure to recognize the octave relationship between "A" and "H" in the stimulus set.
subjects and that results suggest that further training with feedback might have produced additional improvement.

The validity of this study is suspect for several reasons. First, it is doubtful that meaningful results regarding the effects of training could arise from a 2-week study. Not only was the training regimen completed in six sessions but the effect of training with feedback was based on measurements taken after only one treatment. Furthermore, the amount of the improvement was quite small: 2.17 bits of information transmitted by the subjects is roughly equivalent of the correct judgment of four to five tones, while 2.38 bits is just slightly greater than the correct judgment of five tones. Moreover, the choice of stimuli was an exceedingly poor one, especially with respect to the frequencies chosen. The asymmetrical relationship of the semitones between the notes A, B, C, F, and G forms a meaningful context in several keys, most notably A Minor. Additionally, the stimuli were unbalanced with respect to A, because “H” was merely a duplication of A at the octave. Furthermore, it is possible, and quite probable, that Vianello’s subjects determined the identities of these tones through the use of relative pitch cues even though they were nonmusicians.

Cuddy (1968)

In 1968, Canadian psychologist Lola L. Cuddy published an account of her experimentation on absolute pitch. This account was largely drawn from her 1965 dissertation. Cuddy had investigated the effects of training, especially training on just one tone, but also had attempted to resolve the discrepancy between the small number of pitches that average listeners could
correctly identify and the large number that possessors of absolute pitch could identify.

Cuddy’s first experiment examined the relationship between the level of listeners’ musical training and their absolute pitch naming ability. The experiments of Riker (1946) and Oakes (1951) had found that music students could identify the pitches of piano tones far more accurately than untrained listeners, and Cuddy wished to expand upon these studies using sinusoidal stimulus tones. Cuddy’s first experiment used a total of 88 subjects. Her training subjects, 14 males and 14 females, were volunteer undergraduate music students at the University of Toronto. All were studying either vocal or instrument music and were currently enrolled in, or had completed, an aural skills curriculum. Her general listeners, 20 males and 40 females, were volunteers from arts courses at the university and, although none was majoring in music, some had had musical training. Various 12-tone series were played on a tape recorder through a loudspeaker located 4 feet from the subjects at a sound level of approximately 75 dB. Six 12-tone series were prepared, each in a unique order. Four of the series were realized by a sine tone generator and the other two by an experienced pianist on a recently tuned (A4 = 440 Hz) piano. The frequencies of the tones were restricted to those used in standard equal temperament, and tones were randomly selected, four each from each third of a pitch range spanning C#2 to C7. The frequencies and notation of the selected tones are shown in Figure 3.
Figure 3. The musical notation, note names, and frequencies in Hertz for the 12 pitches used in Cuddy’s Experiment 1 (1968).

While these 12 notes were quasi-randomly selected from the 60 possible notes within Cuddy’s specified range, they comprised a skewed sample. While the pitch classes D, F#, and A# were not represented at all, C was represented twice and A was represented three times. Each 2-second tone had an interstimulus time interval of 10 seconds. Subjects were presented all six series and instructed to write the note name and octave designation of each tone. A cardboard simulation of a piano keyboard marked with the note names and the USA Standard octave designations was provided as an aid. No correction was provided during this experiment and subjects were tested either individually or in small groups.

Cuddy concluded that either recent learning or previous experience had a substantial influence in subjects’ absolute pitch naming ability. While there was a substantial amount of overlap between the musical and general listeners, in general, the music students could judge pitch more accurately than the nonmusic students. For the sake of analysis, Cuddy separated the scores of the 28 music majors into two groups—those of the 16 piano
students, and those of the 12 nonpiano students. This enabled Cuddy to conclude that "the judgment of sine-wave tones appears to be related to musical training, while the judgment of piano tones appears to be related to piano training." (Cuddy, 1968, p. 1071) Cuddy also reported that 6 of the pianist, music majors and 2 of the nonmusic majors (a former singer and a former pianist) exhibited an average error of less than one semitone on the piano tone portion of the test. She used this criterion as the maximum acceptable level of error to justify the designation of absolute pitch. Believing that this experiment confirmed the relationship between musical training and absolute pitch judgment, Cuddy designed a training experiment to determine whether or not a causal relationship could be verified.

Cuddy, who cited Paul Hindemith's *Elementary Training for Musicians* among her influences, developed a training technique designed to instill in her subjects a single internal standard, A4. Six female subjects who did not meet the absolute pitch criterion in Experiment 1 were selected from the group to participate in a training experiment. Three, who were music majors, already held diplomas in music performance and had an average of 13.3 years of musical training. The other 3 subjects were nonmusic majors who had received an average of 5.7 years of musical training. These subjects were trained (and tested) using only sine tones, pitched in the 12 equal-tempered semitones from E4 to D#5. Cuddy constructed four levels of training tapes, graduated according to difficulty. Each level contained four tapes of 24 randomly ordered tones in which the representations of A4 steadily decreased. The tone A4 made up one-half of the tones in Level I, three-eighths of the tones in Level II, one-fourth in Level III, and one-eighth in Level IV. At all levels, the tones other than A4 were randomly selected from
the 11 remaining tones in Cuddy's single octave range. A 5-second duration was allotted to both the training tones and the interstimulus time interval.

Subjects were trained and tested individually, three or four times each week and testing sessions preceded each training session. All subjects began at Level I and progressed to the next level when their scores reached an accuracy of no more than two errors. The feedback procedure was unusual. Subjects were instructed to call out "A" within 2 seconds of a tone's onset, whenever they judged the tone to have that pitch. The response had to be quick because the sounding of an A4 on the training tape caused a light signal on top of the loudspeaker to activate 2 seconds into its duration, and to stay lit for the remaining 3 seconds. Table 6 shows the effect of feedback in the four possible training scenarios.

Table 6. The four possible training scenarios in Cuddy's Experiment 2 (1968).

<table>
<thead>
<tr>
<th>Scenario One</th>
<th>Scenario Two</th>
<th>Scenario Three</th>
<th>Scenario Four</th>
</tr>
</thead>
<tbody>
<tr>
<td>A4 is sounded</td>
<td>A4 is sounded</td>
<td>Non-A4 is sounded</td>
<td>Non-A4 is sounded</td>
</tr>
<tr>
<td>Subject calls &quot;A&quot;</td>
<td>Subject is silent</td>
<td>Subject calls &quot;A&quot;</td>
<td>Subject is silent</td>
</tr>
<tr>
<td>Light comes on</td>
<td>Light comes on</td>
<td>Light stays off</td>
<td>Light stays off</td>
</tr>
<tr>
<td>Correct Judgment</td>
<td>Omission Error</td>
<td>False Alarm Error</td>
<td>Correct Judgment</td>
</tr>
</tbody>
</table>

Following a pretest, the 3 music majors practiced with all four training tapes of their level (randomly selected) and were tested after each session. Their final test consisted of two tapes randomly selected from among six test tapes. The 3 nonmusic majors followed a different pattern. They were given two separate pretests, four training sessions, a test session, four more training
sessions and a final test session.\textsuperscript{28} Surprisingly, they were allowed to hear an A4 for 5 seconds before the test began.

For testing purposes, Cuddy constructed six new tapes of 120 tones each. Each tape contained 10 presentations of each of the 12 tones in a quasi-random order (no tone could occur more than twice in succession). Each test tone was presented for 1 second and the interstimulus time interval was 5 seconds. During testing, subjects responded by writing an “A” or an “X”; no feedback was provided. Table 7 shows each test item’s four possible outcomes.

Table 7. The four possible testing scenarios in Cuddy’s Experiment 2 (1968).

<table>
<thead>
<tr>
<th>Scenario One</th>
<th>Scenario Two</th>
<th>Scenario Three</th>
<th>Scenario Four</th>
</tr>
</thead>
<tbody>
<tr>
<td>A4 is sounded</td>
<td>A4 is sounded</td>
<td>Non-A4 is sounded</td>
<td>Non-A4 is sounded</td>
</tr>
<tr>
<td>Subject writes “A”</td>
<td>Subject writes “X”</td>
<td>Subject writes “A”</td>
<td>Subject writes “X”</td>
</tr>
<tr>
<td>Correct Judgment</td>
<td>Omission error</td>
<td>False Alarm Error</td>
<td>Correct Judgment</td>
</tr>
</tbody>
</table>

Cuddy stated that:

Test performance was scored by dividing the number of correct identifications of A by the total number of A’s called, and multiplying the quotient by 100. This score may be denoted 100R/(R+W), where R is RIGHT identification when A is presented, and W is WRONG, or false alarm. (Cuddy, 1968, pp. 1072-1073)

Note that this measure is a far cry from either the subject’s percentage of correct responses made to A4, the true issue, or, the subject’s percentage of correct responses to all tones—the score most commonly reported in measures of absolute pitch. Cuddy’s measure reflects neither the omission errors of Scenario Two, nor the correct judgments of Scenario Four. To

\textsuperscript{28}Cuddy provided no explanation as to why the two groups were treated differently, or what effect this might have had on the results of the training.
achieve a score of 100 a subject merely had to avoid false alarm errors, and to respond correctly to at least one A4. Nothing else bore any consequence.

Cuddy reported that all 6 subjects improved after training and 2 of the 3 music majors achieved perfect scores on their tests of absolute identifications of A. One music major, whose pretest score was less than 20, achieved a 100 by session four and remained at this level until session eight, at which time testing was discontinued. Another, whose pretest score was 40, achieved a score of 100 by session 12, her final session. The third music major, whose pretest score was approximately 10, achieved only a 60 after her 16th and final session. Her personal best was a score of 70, occurring on session 11. The nonmusic majors (who were tested only after their fourth and eighth training sessions) also showed improvement. While none ever scored higher than 50, two of the three showed improvement over the pretest after the first test, and all showed improvement from the first test to the second. After concluding that subjects can improve in their ability to identify A4, Cuddy designed an experiment to see if this A4 training would be effective in improving subjects’ abilities to identify a series of tones.

Cuddy’s third experiment utilized 10 music majors. Five were randomly assigned to an experimental group that trained on the identification of A4 in the manner of Experiment 2. The other 5 were assigned to an experimental group that trained on all tones equally represented with correct identifications provided after every response. In her attempt to study the effect of A4-weighted training on absolute pitch, Cuddy (perhaps for mathematical simplicity) chose to use 10 rather than 12 different chromata as stimuli. Four test tapes presented 10 presentations of each of the 10 semitones from F4 to D5, inclusive. The tones were randomized, except
that no tone could occur no more than twice in succession. The sine tone stimuli were 1 second in duration and the interstimulus time interval was 10 seconds.

Four tapes were used by the group that trained with an equal representation of all tones. The training tapes were identical to the 40-tone test tapes, except that each contained different random orders. Correct identifications were provided after every response during training. Only 10% of the tones on these tapes were $A_4$.

For the group that received the $A_4$-training, Cuddy prepared four levels of training tapes of graduated degrees of difficulty. Each level contained two tapes of 40 randomly ordered tones with decreasing representations of $A_4$. Of the tones in the Level I training tapes 40% were $A_4$; 30% were $A_4$ in Level II, 20% in Level III, and 10% in Level IV. The duration of the training tones was shortened from 5 seconds (as in Experiment Two) to 1 second, in response to suggestions from participants in the previous experiment. The interstimulus time interval was lengthened to 10 seconds; it had been 5 seconds in the previous experiment. Whenever the 1-second stimulus tone was $A_4$, 5 seconds after its onset (the response period) the training light came on and $A_4$ was sounded for another second, regardless of the response. No feedback was provided for any other tones. Table 8 shows the four possible training events and the feedback available to the subjects.
Table 8. The four possible training scenarios in Cuddy’s Experiment 3 (1968).

<table>
<thead>
<tr>
<th>Scenario One</th>
<th>Scenario Two</th>
<th>Scenario Three</th>
<th>Scenario Four</th>
</tr>
</thead>
<tbody>
<tr>
<td>A4 is sounded</td>
<td>A4 is sounded</td>
<td>Non-A4 is sounded</td>
<td>Non-A4 is sounded</td>
</tr>
<tr>
<td>Subject calls “A”</td>
<td>Subject is silent</td>
<td>Subject calls “A”</td>
<td>Subject is silent</td>
</tr>
<tr>
<td>Light comes on</td>
<td>Light comes on</td>
<td>Light stays off</td>
<td>Light stays off</td>
</tr>
<tr>
<td>A4 is sounded</td>
<td>A4 is sounded</td>
<td>Nothing sounded</td>
<td>Nothing sounded</td>
</tr>
<tr>
<td>Correct Judgment</td>
<td>Omission Error</td>
<td>False Alarm Error</td>
<td>Correct Judgment</td>
</tr>
</tbody>
</table>

Both experimental groups underwent eight training sessions between the pretest and posttest. During training, each session consisted of practice with one of the training tapes appropriate for their group. Before all testing, both groups were, for some undisclosed reason, given one complete hearing of all 10 stimulus tones sounded in ascending order.

By comparing the mean number of correct judgments for each of the 10 tones on the posttest and pretest, Cuddy found that the effect of training was negligible for the group that trained with equal probabilities of A4. The other group, however, was able to increase the mean number of correct responses to each of the 10 tones, and frequently by large measure. Cuddy performed an analysis of variance measure to verify that the difference in the pretest and posttest scores was dependent on the method of training.

In summary, Cuddy’s first experiment suggested that specific training and past experience play an important role in pitch naming ability. After her second experiment she concluded that pitch naming can be improved through training on the identification of just one tone. Her third experiment showed that this training method was more effective than ordinary practice with correction. However, one can certainly criticize this study for its lack of
musical stimuli in the training experiments, its relatively long interstimulus
time intervals (which are certainly conducive to the use of relative pitch), and
its presentation of the tones, whose storage in long-term memory is at issue,
immediately before testing. As far as the applicability of the results to the
issue of absolute pitch is concerned, only Experiment Two utilized all 12
chromata, but here the response task was one of identifying “A” or “not A.”
The only experiment that involved pitch naming omitted D# and E. Cuddy
rightly expressed the concern that “A training encourages the listener to adopt
a strategy of subvocalizing the tone A and of maintaining this subvocal
response throughout the testing session.” (Cuddy, 1968, p. 1075) This is a
serious consideration, for, while subvocalizing may lead to increased accuracy
in various pitch identification tasks, its inherently relative nature has little or
no bearing on what is considered by most to be absolute pitch skill.

Brady (1970)

Probably the most remarkable training experiment in the absolute pitch
literature was reported by Paul T. Brady (1970a, 1970b), who proclaimed
himself to be the first previously unskilled adult to achieve, on his own
initiative, a level of nearly perfect semitone identification of pitches.29 Brady,
an active amateur musician who worked at Bell Telephone Laboratories,
reported, in anecdotal fashion, his own experiences and training techniques.

29Brady’s statement refers to the findings of W. Dixon Ward’s (1963) literature
review. Ward wrote “the consensus of the experiments on learning is that
‘genuine’ AP [absolute pitch] cannot be taught to adults. Although some
improvement with training will occur, no one has yet brought an initially
unskilled subject to the level of proficiency possessed by a fair number of
musicians, i.e., nearly-perfect semitone discrimination.” (p. 39)
In 1969, Brady renewed his interest in attempting to acquire absolute pitch, after having been unsuccessful on two previous occasions. In 1958, he had failed to memorize an A on the piano by playing it frequently, and a similar attempt a few years later also ended in failure. Brady had played the piano since he was 7 years old and sung in choruses since he was 12. He also possessed some experience at harpsichord tuning. Yet, Brady (1970a) stated:

In short, I have had plenty of musical performing and listening experience, much of it in tasks requiring tone discrimination. Not once was I ever aware of the key of a played piece nor could I identify a note. It would have been an arbitrary guess.\(^{30}\) (p. 884)

In February 1969, at age 32, Brady began a 3-month training program to try to discover any attributes of pitch that might have previously escaped him. He “spent hours thinking of notes, mentally playing pieces, and trying to ‘hear’ previously undetected differences—the new dimension. I also tried playing piano pieces in foreign keys to see if they sounded different. They didn’t.” (p. 884) Brady mentioned that the only positive result of this program, which also included practice with various musical texts, was an improvement in his relative pitch.

In May 1969, Brady proclaimed that a “breakthrough” occurred immediately after he implemented a different technique, which he adapted from the second and third (A\(_4\)-weighted) experiments of Cuddy (1968). Brady listened to tapes of computer-generated sine tones that contained decreasing proportions of C. The sine tones on his tapes spanned the three-octave pitch

\(^{30}\text{Brady’s comments here refer to “absolute tonality” as well as absolute pitch. It is assumed that any process used in the aural absolute determination of a musical composition’s key would, at some point, involve the determination of its tonic. This ability, not necessarily related to absolute pitch, does not fall within the scope of this dissertation.}\)
range from A#2 (117 Hz) to A5 (880 Hz). Brady's use of the pitch class C, as opposed to Cuddy's use of A4, was apparently an arbitrary decision. It is important to note, however, that while Cuddy weighted her tapes heavily with an A4, Brady let his C-weighting be randomly selected from the three C's within his designated pitch range (C3, C4, and C5). Lists of the tones were printed by a computer that also randomized them in a manner which permitted tones of the same pitch class, but not the same frequency, to repeat themselves. Equal distributions of tones other than C's made up the remainder of each tape. Brady produced 10 hours of tapes in which the probability of C occurring dropped from 0.4 to 0.083 (1 out of 12, or chance level). The interstimulus time interval on these tapes also decreased from 5 seconds to 1.4 seconds. The duration of the tones was held constant at 1 second and all tones were played through a loudspeaker at equal voltage levels.

Brady's training technique also differed from Cuddy's in other details. Brady tried to identify the pitch class name of every tone before looking it up on the list, but throughout this process he endeavored to remember C5 (523 Hz). He avoided humming, consciously tried not to use relative pitch, and tried to judge each tone in terms of harmonic scale-steps in a C tonality.

Brady (1970a) wrote:

One never identifies notes from "interval from previous note" techniques, unless the previous note is C. For example, the sequence G-A should not sound like a whole step; the G should sound like C's dominant, and A like C's minor. (p. 884)

---

31 The inclusion of all 12 pitch classes and a range greater than an octave are both significant improvements over Cuddy's (1968) methodology.
After training with this technique for 1/2 an hour daily for 2 months (approximately 60 hours total) Brady gradually began to retain the unique chroma of each pitch class. He reported that, oddly, B♭ was the easiest to remember, followed next by C. He concluded that, by relating the chromata to a C tonality, the task of naming the pitch as on the training tapes was made very easy. Not only was Brady able to identify correctly nearly every note of the equal-probability, fastest-rate tape without feedback, but he could even do it while reading a book. He also stated that not only did his ability to identify sine tones readily transfer to piano tones, but also to other timbral sources such as an automobile horn, telephone dial tone, power appliance hum, and toy bell.

While there was no documented pretest of his abilities before this training, Brady sought to give himself the most stringent piano-tone posttest he could devise. Brady had his wife play a single tone on the piano, just after he awoke, for him to identify. The tones were determined by a list of computer-randomized pitch classes, but she was free to choose the octave. Brady identified only the pitch class name without attempting the octave identification. Testing occurred on 57 consecutive days, concluding on 30 September 1969. On that date, Brady made 37 correct judgments (65%), 18 semitone errors (31.5%), and 2 whole-tone errors (3.5%). Nine of his

---

32 Brady held some notions regarding chroma which are far from the standard use of the term by musicians. Brady wrote: “‘Chroma’ is used to describe the different perceptual sounds that different notes acquire when placed in a scale. For example, the tonic note (C of a C-major scale) has a ‘finality’ about it; it is the accepted last note of a composition. Chroma may also be thought of as that quality which makes all C’s (or D’s, etc.) sound similar.” (p. 887) Here, aspects of this first definition bear resemblance to the modern concept of Stufe, or harmonic scale-step, advanced in the writings of noted German music theorist Heinrich Schenker (1906, trans. 1954).

33 This test procedure is quite similar to one employed by Petran (1932).
semitone errors were guessed too high, nine were guessed too low, and both whole tone errors were guessed too high. The absolute mean error for Brady’s judgments was 0.39 semitones and he regarded his ability to identify correctly 96.5% of the test tones within plus or minus one semitone to meet the often-cited near-perfect semitone identification standard.

Seeking to verify his techniques, Brady trained 2 other subjects for several weeks. One was a female nonmusician. Because she lacked relative pitch skills, her inability to relate to tones in the context of a C major tonality prohibited her from making any progress. The other subject was a female French hornist, majoring in music in college. While she experienced no initial difficulty using the training tapes, she also failed to make significant progress.

In an effort to determine the retention effects of his newly acquired skill, Brady again measured his pitch identifying ability on 2 March 1970, 6 months following the conclusion of his lengthy posttest and more than 8 months after his last training session. Five randomly selected tones were performed on a flute in a soundproof booth. Tones were interspersed with 3 minutes of conversation in an effort to interfere with the memory of the preceding tone. Brady identified four of the tones perfectly and erred on the other by a semitone. He (1970a) wrote “perhaps AP [absolute pitch] is hard to forget, as well as hard to acquire.” (p. 887)

As an epilogue to Brady’s self-experimentation, it should be mentioned that his absolute pitch ability was tested in the course of research conducted by John B. Carroll, Senior Research Psychologist at Educational Testing Service. Carroll (1975) examined the differences in accuracy and decision times in pitch judgment tasks for 9 subjects; 4 who were nonpossessors, 4 who claimed
absolute pitch ability since childhood, and Brady, now 35 years of age, who claimed to have acquired absolute pitch through self-training. Carroll advanced the two-factor theory of absolute pitch whereby nonpossessors use relative tone height and possessors use this as well as chroma standards stored from long-term memory. Carroll also divorced the judgment of chroma from that of octave, derived separate informational measures for their identification, and attributed subject differences in terms of information transmission capacity on either one or two channels. In his summary, Carroll (1975) wrote "the performance of the trained AP subject [Brady] was not distinguishable from that of the remaining AP subjects either in accuracy or decision time." (p. 2) Furthermore, in the course of analyzing his experimental data Carroll noted that:

Throughout this discussion, the data for the TAP [trained absolute pitch] subject [Brady] have been considered along with the data for the EAP [early absolute pitch] subjects to constitute a pool of AP subject data, since the TAP subject's data are such as to make it untenable to regard him as other than a subject drawn from a population of EAP subjects. If we grant this subject's own claim (Brady, 1970) that he did not have AP ability before starting to train himself in it, it appears that AP ability can indeed be trained to a level of accuracy exhibited by persons claiming AP ability from early childhood. (p. 25)

If fact, the only times when Brady's data varied even minutely from that of the other absolute pitch possessors, was when he made perfect identifications that they missed.

Cuddy (1970)

In 1970, 5 years after completing her doctoral dissertation and 2 years after a condensed report of the dissertation study was published as a journal article, Lola Cuddy again published on the effect of training on the absolute
identification of pitch. This experiment, largely based on the third part of her earlier experiment, sought to compare two differing training techniques. She called the first technique “reference training” (analogous to her previous “A training”); the second she called “series training” (as before). Cuddy (1970) wrote:

The success of reference training implies the presence of a cognitive tonal structure with the reference tone as a nodal point. For example, music students, who study the structure of Western tonality (that is, its intervals and scales), may develop a cognitive structure that maps the characteristics of musical structure. (p. 265)

Cuddy’s 1970 experiment sought to measure the effect of an increased number of reference standards and to determine whether the effect of “reference training” could be generalized to nonmusicians.

Cuddy used paid volunteers recruited from psychology classes at Queen’s University, Kingston, Ontario as subjects. They were not pre-selected for musical experience. Subjects were classified and divided into two groups: one made up of those with some musical experience or interest (Group M), the other containing those with none (Group NM). Musical experience or interest was determined solely by their response to the question, “Have you engaged in any musical activities during the past 3 years?” (Cuddy, 1970, p. 265) Twenty-five subjects ranging in age from 18 to 39 began the experiment, but one dropped out. The 24 remaining subjects still maintained the balance, with 12 in each category.

As summarized earlier, 5 music students who were trained to recognize a single reference tone (A4) were able to improve their ability to judge ten sinusoids. Five others who were trained with “series training,” where tones appeared with equal probabilities followed by immediate feedback, did not show significant improvement.

Cuddy’s report was inconsistent, describing the 25 subjects as “15 men and 11 women.” (p. 265)
Cuddy used two sets of nine sinusoidal tones for her stimulus patterns. These tones were spaced according to their pitch distance in mels and thus were unrelated to any musical scale. The tones of Set A were spaced equally every 200 mels, and spanned a range from 400 to 2000 mels. The tones of Set B were not equally spaced, but grouped into three areas of pitch, low, middle, and high. The distance between tones within these groups was 100 mels, while the boundary distance separating the three groups was 400 mels. The total range of Set B was from 500 to 1900 mels, roughly equivalent to that of Set A. Converted to Hertz, both of these sets fall within the range of 290 Hz to 3000 Hz. Figure 4 displays the relative spacing of the two sets.

![Graph showing the relative spacing of Set A and Set B in Cuddy's (1970) Experiment.]

Figure 4. The relative spacing of Set A and Set B in Cuddy's (1970) Experiment.

Cuddy, in summary, had formed a total of four experimental conditions. There were two different types of training procedures, “reference” and “series,” and two different types of stimuli, grouped and ungrouped.
Eight total cells existed as 3 subjects from Group M and 3 from Group MN were randomly assigned to each condition.

All sine tones used in this experiment were produced on a Hewlett-Packard Audio Signal Generator and recorded on a tape recorder. Subjects listened through a pair of monaural headphones at a level which varied randomly over 15 dB at an average level of 70 dB SPL. This was done in order to minimize differential loudness cues.

The 12 subjects who were “series trained” used tapes composed of six occurrences of each of the nine tones (whether they were from Set A or Set B) in random order. Two such tapes of 54 tones were made for each group. The tones of 600 mels, 1200 mels, and 1800 mels were chosen as the references for the 12 subjects who were reference trained. These tones, designated L, M, and H, occupied the second, fifth and seventh positions in both Set A and Set B and were presented in decreasing proportions at each of the four succeeding levels of training. Equal representations of each of the three reference tones made up 42 of the 54 tones in Level I (77.7%), 36 of the 54 tones in Level II (66.6%), 30 of the 54 in Level III (55.5%), and 24 of the 54 in Level IV (44.4%). At all levels, tones other than the three references were randomly selected from the six remaining tones in the set. All training tapes contained double presentations, so that each of the 54 one-second tones was repeated 4 seconds later.

Each subject participated for six separate sessions; the first for pretesting, the middle four for training, and the last for posttesting. The sessions took place on separate days, and each subject’s participation was completed within approximately 2 weeks. Cuddy constructed five additional tapes of both sets for testing purposes. These tapes offered nine random
presentations of the nine stimulus tones, for a total of 81 presentations. The
duration of each tone was 1 second, and the interstimulus time interval was 4
seconds. Three of the five tapes of each set were randomly chosen and
administered as a pretest and three were again chosen and administered as a
posttest. As in her dissertation experiment, Cuddy once again presented all of
the test tones to her subjects immediately before they were to make their
absolute judgments, thereby introducing an extraneous variable, relative
pitch. Subjects were told that the names of the tones were L-, L, L+, M-, M,
M+, H-, H, and H+; then the tones were presented in random order. Subjects
were also informed that during the test equal numbers of each tone would
occur in a random order. A short practice period ensued (30 judgments in
length), allowing the subjects to familiarize themselves with the testing
procedure. Subjects recorded their responses in an answer booklet. No
correction was provided.

Both training and practice testing took place from the second to the
fifth session. Each of these intermediate sessions presented two training tapes
followed by a test tape. Under “series training,” subjects identified each tone
using all nine labels, and feedback was provided by means of a light panel in
front of them. Under “reference training,” subjects attempted to identify each
tone by indicating whether it was one of the three references and, if so,
whether it was L, M, or H. Subjects using “reference training” were permitted
to advance to the next level when 92.5% of their judgments were correct (100
of 108).

Cuddy (1970) wrote that the use of three separate analytic measures on
the test data “failed to reveal any differences in performance attributable to
stimulus spacing.” (p. 266) This finding is in general agreement with the
grouping experimentation of Pollack (1952). Following this observation, Cuddy reported only data which she had averaged across both stimulus sets (A and B). Table 9 summarizes Cuddy's averaged pretest and posttest results.

**Table 9.** The percentage of correct judgments and bits of information transmitted, averaged across both Set A and Set B, for Cuddy's (1970) Experiment.

<table>
<thead>
<tr>
<th>Correct Judgments</th>
<th>Information Transmitted</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pretest</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Group M Reference</td>
<td>39.0%</td>
</tr>
<tr>
<td></td>
<td>1.31 bits</td>
</tr>
<tr>
<td>Group M Series</td>
<td>37.2%</td>
</tr>
<tr>
<td></td>
<td>1.26 bits</td>
</tr>
<tr>
<td>Group NM Reference</td>
<td>34.0%</td>
</tr>
<tr>
<td></td>
<td>1.18 bits</td>
</tr>
<tr>
<td>Group NM Series</td>
<td>32.0%</td>
</tr>
<tr>
<td></td>
<td>1.18 bits</td>
</tr>
</tbody>
</table>

Cuddy found that the difference between her two training techniques was related to the subjects' musical experience or interest. Group M subjects improved more from training based on the identification of the three reference standards, while Group NM subjects derived more benefit from the equal-probability "series training." Subjects from both groups improved following training, but in all cases Group M subjects outperformed those of Group NM. The highest level of ability displayed by any group of subjects was that of Group M after receiving "reference training."

In summary, it is not hard to imagine why Ward & Burns (1982) referred to this study as "the nonmusical experiment." (p. 266) The use of
nine equally and unequally mel-spaced sinusoids makes faint, if any, reference to the ability of a musician with absolute pitch to identify musical stimuli with near-perfect semitone accuracy. Also, the classification of psychology students as being musical or not on the basis of their response to a single question is highly suspect. The evaluation of the merit of a training technique after a training period of only four sessions, and the use of such a small number of subjects (3 in each cell) are disturbing as well. Even more troublesome is the fact that the subjects heard the tones they were about to “absolutely” judge immediately before testing began. Lastly, little faith can be put in an improvement rate as meager as that displayed here. The group of psychology students with musical experience or interest that was “reference trained,” Cuddy’s group that improved the most, transmitted only 1.91 bits of information following training. This level is still less than the perfect identification of only four tones.36

Cuddy (1971)

In 1971, Lola Cuddy published her third and final experimental article on the topic of acquiring absolute pitch through training. In this article (1971), she investigated the effects of manipulating both the interstimulus time interval and the pitch spacing. Cuddy was especially interested in the effect of musical versus unmusical pitch spacings on learning. She expressed concern that while the mathematical rules for determining the frequencies used in 12-tone equal temperament suggest a two-dimensional system (the octave ratio

36Recall that Pollack (1952) found that subjects could transfer between 2.2 to 2.3 bits of information when judging, through headphones, sine tones spaced at equal logarithmic intervals without any training whatsoever. (The number required for the perfect identification of five tones is 2.32 bits).
is 2:1, and the 12 tones within an octave form equal logarithmic intervals), “this mathematical analysis does not necessarily mean that the perceptual system for pitch recognition is two-dimensional.” (p. 43) Citing work that denies the octave’s value as a basic unit of melodic pitch distance (a mel scale study), Cuddy designed three experiments to determine the influence of octaves among pitch sets to be learned.

Cuddy used volunteers from Queen’s University and a local high school as subjects. All were “music students” who were studying privately. All had some previous keyboard training and some were studying an instrument other than piano as well. All scored at or above Grade VIII level at the Royal Conservatory of Music at Toronto. Their ages ranged from 14 to 38 years.

Cuddy designed three different stimulus sets to be evaluated. Each set contained 12 sine tones of which the 1st, 5th and 12th tones were common to all three sets. Set T represented Cuddy’s triad spacing and comprised four ascending root position F Major triads (F3, A3, C4, F4, A4, C5, F5, A5, C6, F6, A6, and C7). Set KI represented Cuddy’s keyboard interval spacing and comprised a single representation of each of the 12 chromata distributed through the range from F3 to C7 (F3, A#3, D4, F#4, A4, C#5, E5, G5, B5, D#6, G#6, and C7). Tones were separated by 3, 4, or 5 semitones. Set AI represented Cuddy’s arithmetic interval spacing and—other than the first, fifth, and twelfth tones—was not based on the normal equal tempered A4=440 Hz tuning system. Cuddy arbitrarily chose to divide the pitch range 175 Hz to 440 Hz (F3 to A4) into four equal arithmetic intervals and the pitch range 440 Hz to 2093 Hz (A4 to C7) into seven equal arithmetic intervals. The resulting 12 frequencies in the AI set were 175, 242, 308, 374, 440, 676, 912, 1148, 1384, 1620,
1856, and 2093 Hz. All sine tones were generated on a Hewlett-Packard function generator, recorded on magnetic tape, and played with a Crown tape recorder. Subjects listened monaurally through a pair of high quality headphones at a level of approximately 70 dB ± 5 dB. This random variation was allowed to minimize possible differential loudness cues.

The response task was the same for all of the three following experiments. Subjects were given a booklet, each page of which contained 12 columns of integers from 1-12. Rows represented the 12-tone sets, and columns represented trials. The response task was to circle the number of the tone presented. Cards identifying the construction, note names, and frequencies of each set were available during testing.

Test tapes for Experiment 1 and Experiment 3 provided three randomly ordered presentations of the 12 tones in a set. Four such tapes were constructed for each of the three stimulus sets. Stimulus tone duration was 1 second and the interstimulus time interval was 4 seconds. Eight similar test tapes were constructed for Experiment 2: four with 3-second interstimulus time intervals, and four with 8-second interstimulus time intervals.

Experiment 1 was designed to measure the effect of the three spacings on subjects' abilities to judge tones. It was not a training experiment. Twelve subjects from the previously described subject pool were asked to identify four different random orderings of the three different sets (T, KI and AI). Before the test began, subjects were presented, in ascending order, all 12 tones of the set to be judged. They were also given time to run through at least 24 practice trials. No feedback was provided during testing. The music students judged Set T with greater accuracy than either Set KI or Set AI. The mean
percentages of correct responses for Sets T, KI and AI were 60.8%, 47.7% and 40.8%, respectively.

Experiment 2 was designed to further explore differences in subjects’ abilities to identify Sets T and KI. This experiment attempted to determine the relationship between accuracy and level of musical experience, and the effect of varying the interstimulus time interval. Seven music students who had completed Grade X or higher were placed in Group 1, and seven others who had completed only Grades VII or VIII were placed in Group 2. Four college-age subjects who were not studying music were tested (months later) to serve as a control group. Each group judged tones in four experimental conditions, using Set T and Set KI with interstimulus time intervals of 3 and 8 seconds. The testing procedure was similar to that used in Experiment 1. Cuddy found that, averaged across all conditions, the subjects in Group 1 outperformed those in Group 2. Their mean percentages of correct responses were 66.8% and 42.3%, respectively. The control subjects scored significantly lower. Both groups of music students identified tones in Set T more accurately than in Set KI, and across all conditions their mean scores for the 3- and 8-second interstimulus time intervals were nearly identical. Neither differences among sets nor rate of presentation produced significant difference in the accuracy of responses by control subjects.

Experiment 3 was the training experiment. Fifteen music students were randomly divided into three groups. Each 5-subject group was then assigned either Set T, Set KI or Set AI. They were given a pretest, eight training sessions, and a posttest, each limited to the appropriate set. The testing procedure was identical to that of Experiment 1. Cuddy constructed A4-weighted training tapes similar to those used in her earlier
experimentation. She constructed eight tapes for each of the three stimulus sets. Two tapes were made at each of four levels, and each level contained decreasing proportions of A4. The A4 reference tone made up 9 of the 24 tones in Level I (37.5%), 6 of 24 in Level II (25%), 3 of 24 in Level III (12.5%), and 2 of 24 in Level IV (8.33%). At all levels, tones other than the A4 reference tone were randomly selected from the 11 remaining tones in the set. The sessions consisted of practice on two training tapes, and subjects progressed to the next level when they completed both tapes with no more than three errors. All sessions were scheduled at least 24 hours apart.

Stimulus tone duration was 1 second and the interstimulus time interval was 4 seconds. During training subjects were instructed only to judge tones as "A4" or "not A4," and were provided feedback only on those tones that were A4.

Cuddy did not provide individual pretest and posttest scores, but instead provided the percentage of average improvement from the pretest to the posttest—i.e., gain scores.37 She reported that the average relative improvement for the Set T group was 68.4%, and accuracy levels ranged from 42% to 85%. The average relative improvement for the Set KI group was 20% in a range from 0% to 50%. Finally, the average relative improvement for the Set AI group was 24% in a range from 9% to 55%. Cuddy stated, in reference to the Set T group:

For three of the five listeners, the only errors made at the post-test were at distance 131 [the octave] and these errors comprised only 6 per cent of all responses made to the post-test by the three listeners. (p. 52)

37The pitfalls of relying on gain scores are discussed in Chapter 5.
If this was the case, her following statement, "in terms of information transmitted (T), three listeners achieved the maximum possible value of 3.59 bits with Triad spacing" (p. 53) must be incorrect. While it is impossible to have committed errors and still have an informational score equivalent to the perfect identification of all 12 tones (unless these octave errors were now suddenly regarded as correct responses), we should not overlook the fact these musicians did well, extremely well, on the posttest.

No data were provided for the other two groups that had not first been averaged among them. Probably the closest measure of actual training of absolute pitch by musicians in this experiment would have been found in the results obtained by those practicing with Set KI. These were unfortunately lost when Cuddy combined them with the results of the subjects practicing with Set AI. Yet, Cuddy wrote:

The experiments indicate that a set of tones based on the F-A-C Triad and its octaves are judged more accurately, and learned more rapidly by music students than sets of tones without simple chord or octave relationships. (p. 53)

Cuddy apparently believed that her subjects achieved an unusual form of absolute pitch, one that enabled them to identify absolutely tones from a major triad dispersed in four octaves. A more likely interpretation is that the training merely refined their relative pitch ability (used in determining the root, third and fifth of the triad) or their octave-judging ability, and that no absolute judgments even occurred. In fact, it is hard to imagine how trained musicians could score low enough on the pretest to allow such a high gain score, or how any errors at all could be made on a posttest of this design following practice.
Heller & Auerbach (1972)

In 1972, Morton Heller and Carl Auerbach of Yeshiva University conducted an experiment in an attempt to replicate Cuddy’s 1968 findings. They found it surprising that A4-weighted training (where subjects judge only whether the presentation of a tone is or is not an A4, and receive feedback only when the tone presented actually was A4) produced better results than training programs in which subjects identified all stimulus tones and received complete feedback. In fact, Cuddy’s regular (“series”) training group did not improve at all. Heller and Auerbach suspected that the subjects in the AFB group (A-feedback) received more training than the RFB group (regular feedback). They surmised that the AFB group actually heard more A’s than the RFB group, and were trained to a criterion of two errors or fewer before advancing to the next level.38

Heller and Auerbach took great care to replicate the experimental conditions of Cuddy’s (1968) Experiment 3. They used 10 sine tones from the standard equal tempered (A4= 440 Hz) scale as stimuli. The tones ranged from F4 (349 Hz) to D5 (587 Hz). The duration of each tone was 1 second and each interstimulus time interval was 10 seconds. The stimulus tones were tape recorded and presented to subjects through headphones. Heller and Auerbach did not, however, devote as much care as Cuddy had to the selection of their subjects. Each of Cuddy’s groups comprised 5 music majors. For their replication, Heller and Auerbach (1972) wrote, “each group consisted

38Cuddy’s response to these issues, as well as others, can be found in Cuddy (1972). She stated that while the subjects in the A-weighted training did receive a greater number of A’s to judge, they did so by design. This was the primary intent of this specific training regimen. However, she argued that the total number of judgments made during training was the same for both groups.
of eight Ss who had been recruited from a student population and randomly assigned to one of the two groups. The Ss had no special musical training." (p. 222)

Both experimental groups, RFB and AFB, followed the same schedule of events. First, a period of familiarization was allowed. Subjects were presented the 10 tones in ascending order and were told their correct letter names. Second, a pretest of eight presentations of all 10 tones was given. Orders of tones were quasi-random; no tone could be "repeated [italics added] more than twice in succession." (p. 222) Subjects had to identify all 80 tones without the benefit of feedback. Third, both groups of subjects trained on four tapes, each with a decreasing proportion of A's. This design differs significantly from Cuddy's. Here, both groups used the same tapes in which 40% of the tones in Level I were A4, 30% were A4 in Level II, 20% in Level III, and 10% in Level IV. During training, subjects in the RFB group were asked to identify, in writing, each of the 1-second tones. Five seconds into each 10-second interstimulus time interval, a recorded voice provided the correct response for the item heard just before. During training, subjects in the AFB group were also asked to identify, in writing, each of the 1-second tones, but they had only to state whether they judged that the tone was an A. Five seconds into the 10-second interstimulus time interval, a recorded voice stated "A" if the tone belonged to pitch class A, or nothing if it did not. Lastly, a posttest was given which differed from the pretest only in the order of the stimuli. Heller and Auerbach stated that the entire experiment was completed in 2 days. On day one, subjects were pretested and subsequently completed training on the first two levels. On day two, subjects were posttested following completion of the remaining two levels.
Heller and Auerbach stated that the initial level of performance was equivalent for both groups and, furthermore, that both groups improved equally with practice. The pretest and posttest mean percentages of correct responses for the AFB group were 27% and 37%, compared to 27% and 38% for the RFB group. The mean judgment error in semitones on the pretest and posttest fell from 1.44 to 1.00 for the AFB group, and from 1.40 to 0.92 for the RFB group. Heller and Auerbach suggested that Cuddy's experimental findings (that A4-weighted training was effective and regular training was ineffective) were not the result of differences in feedback technique, but the result of "differences in the stimuli to which the groups were exposed." (p. 223) They further suggested that improved performance by individual members of Cuddy's "series trained" group was most likely hidden when Cuddy averaged the data. In 1972, Cuddy defended her experimental technique and her conclusions. She stated that while 2 of the 5 subjects in her experiment did show improvement following "series training," their group mean percentage of correct responses rose only from 41% on the pretest to 43.5% on the posttest, while the "A-standard training" group rose from 39.5% to 80.5%. Cuddy noted that her subjects were highly trained musicians while Heller and Auerbach's were not. In fact, Cuddy's subjects, music majors, scored higher accuracy levels on their pretests than Heller and Auerbach's subjects scored on their posttests.

One likely interpretation of this polemic is that Cuddy's subjects were able to utilize their relative pitch abilities in the course of training and, unfortunately, in testing as well. On the other hand, Heller and Auerbach's subjects presumably lacked such skills and, as a result, were not able to derive as much benefit from "A-standard training." Of course, little faith can be
placed in the results of a training experiment that attempted to teach a skill as complex as absolute pitch in only 2 days. Improvement as small as that exhibited here, from 27% correct on the pretest to 37% or 38% on the posttest, is hardly convincing evidence that Heller and Auerbach’s subjects had acquired any proficiency at all. Perhaps they were just beginning to acquire some relative pitch skill. In the final analysis, neither Cuddy’s nor Heller and Auerbach’s experiment demonstrates the acquisition of true absolute pitch through training.

**Falk (1975)**

In 1975, French composer Julien Falk published his method of absolute pitch instruction entitled, *Formation Rationnelle de l’Oreille Musicale: Acquisition de l’Oreille Absolue—Traité Complet en Cinq Cahiers: Exercises, Dictées correspondantes, Solfège accompagné* [Rational Formation of the Musical Ear: Acquisition of Absolute Pitch—Complete Treatise in Five Volumes: Exercises, Corresponding Dictations, Accompanied Sight-singing]. Falk’s treatise comprises five volumes, though each volume is only about 30 pages in length, and strives to teach absolute pitch through a combination of solfège and dictation exercises, which range from very easy to impossibly difficult. Falk believed that previous teaching methods had incorrectly focused on having students identify notes after they had been sounded, instead of requiring the student to produce a tone of an indicated pitch. Falk stated that the goal of his training method was to give his students absolute pitch, and he claimed that he had been successful in this endeavor.

The first three volumes of Falk’s treatise contain tonal sight-singing and dictation exercises. The fourth and fifth volumes are dedicated to
atonality, and contain solfège exercises which include piano accompaniments written intentionally to disrupt the singer. Falk stated, in the preface to his treatise, that for the student not be distracted by the accompaniment, he must possess absolute pitch, which the student would have acquired following completion of the third volume. Falk asserted:

Notre affirmation est basée non sur un espoir, mais sur une certitude, car notre enseignement nous a permis d'expérimenter notre méthode sur de nombreux élèves. (Falk, 1975)\(^{39}\)

Each exercise is to be practiced 10 times, in 6 different ways, in the following order. First, the exercise is played on the piano, without naming the notes. Second, the exercise is played on the piano while the notes are named, but not sung. Third, the exercise is sung twice on the neutral syllable "ah," with the help of the piano. Fourth, the exercise is sung twice while naming the notes,\(^{40}\) with the help of the piano. Fifth, the exercise is sung twice on the neutral syllable "ah," without the aid of the piano. Sixth, the exercise is sung twice while naming the notes, without the aid of the piano.

Except for his own testimonial, Falk offers no evidence supporting the claimed success of his training technique. While a method that combines both solfège and dictation exercises might seem meritorious to traditionalists, judgment of Falk's technique should be withheld pending an objective experimental study of its effectiveness.

\(^{39}\)"Our contention is based not on hope, but on certainty, for our teaching experience has allowed us to test our method on numerous students."
\(^{40}\)The names of chromatic notes are to include their chromatic alterations.
Hurni-Schlegel & Lang (1978)

In 1978, Lisbeth Hurni-Schlegel and Alfred Lang, of the Psychologisches Institut der Universität Bern, published an account of their experimentation on absolute pitch. Their account (Hurni-Schlegel & Lang, 1978) detailed four separate experiments. Their account first detailed the development of tests designed to measure absolute and relative pitch abilities. One of two different absolute pitch tests involved a distractor, while a relative pitch test included presentation of a standard tone to be used for comparison. These tests were then administered to two populations. The first group consisted of 451 students representative of the general school population of Bern, Switzerland, and the second group consisted of 80 music majors at the conservatory. The test scores yielded approximately normal distributions from the groups, and moderate to zero correlations between absolute pitch ability and several musical variables.

As part of their experimentation Hurni-Schlegel and Lang conducted a training experiment with children to measure the modifiability of absolute pitch ability. For this purpose they constructed a special training device containing buttons, switches and counters. Twelve buttons were labelled with the pitch names C through B and were arranged as are the keys on a piano. Subjects responded by pressing these buttons. During training, subjects were presented randomly selected square wave tones at the frequencies of the 12 equal-tempered semitones from C4 to B4. Tones sounded uninterruptedly until the subject pressed the correct button. In addition to this immediate feedback, a counter displayed the number of tones correctly judged, thereby enabling the subjects to record the results of each session themselves. In this way all subjects were kept informed of their
progress. Before the beginning of each session subjects could listen to the tone C4. The training program offered two levels of difficulty. The first contained only the seven tones of the C diatonic collection, while the second contained the 12 tones of the C chromatic collection.

The investigators obtained their subjects from a group of children who were attending a week-long extra-curricular music camp. Seven girls and 14 boys, ranging in age from 11 to 15 years, participated in the experiment. All had played a wind instrument for a year, but few had taken private lessons. They were tested on the first and last days of their camp outing. On each of the 5 intervening days the subjects practiced individually for 15 minutes using the training apparatus. During the 5 days of training each subject made a total of 360 training judgments: 50 on the first day, 70 on the second day, and 80 on each of the remaining days. The training regimen was introduced to the children as a competition; those who could achieve either the highest overall score or the greatest amount of improvement would win prerecorded audio cassettes. Consolation prizes were provided for the others. Hurni-Schlegel and Lang reported that this inducement proved unnecessary because the children found the training highly interesting and motivating. A similarly formed group of children who did not train was used as control group.

The absolute pitch tests administered by Hurni-Schlegel and Lang used sine tone stimuli ranging, in equal-tempered semitones, from C4 to C5. The first test followed this sequence: a 1-second stimulus tone, 10 seconds of silence, then the next stimulus tone. The other absolute pitch test followed a longer sequence: a 1-second stimulus tone, 10 seconds of silence, a 4-second
period of audio distraction, another 10 seconds of silence, and then the presentation of the next stimulus tone.

Not all of the 21 children who participated in the learning group could master both levels of difficulty. In fact, in the course of the 5-day training period some never reached the chromatic stimuli. The average percentage of correct responses for the children who remained on the seven diatonic tones increased from 62% to 79.7% following training. The average percentage of correct responses for the 15 children who advanced to the 12 chromatic tones reached only 57% and their progress at this level occurred at a slower rate than that at level one. The authors also reported that, in general, almost all of the children who received training showed some improvement. The posttest results of the combined training group, when compared to either their pretests or the test results of the control group, indicate their significantly greater abilities. Further analyses, which treated half-step errors as correct responses, produced similar findings. The average percentage of correct responses for the entire training group increased from 48.7% on the pretest to 68.8% on the posttest. Two children reached a terminal performance level of 86.6%. This figure is comparable to results of individuals possessing absolute pitch. Hurni-Schlegel and Lang also conducted an analysis of average error-size. Although observable individual differences were established, the mean error as a result of training in the absolute test decreased from 1.75 to 1.25 semitones. Additionally, almost a third of the test subjects' responses were, after training, within the margin of a whole step 90% of the time.

Hurni-Schlegel and Lang’s results indicate that absolute pitch identification in children aged 11-15 can be improved through simple test
repetition within the short training period of 5 days with only 15 minutes of practice daily. The use of a comparatively large subject population and the use—possibly the first—of a control group for comparison of training effects are highly commendable. Less commendable was the brevity of the training period, and the use of nonmusical stimuli of greatly restricted range. Also, unfortunate was the lack of effort to determine what method, if any, the children were using to make their absolute pitch judgments. The disclosure of individual test scores would have allowed the determination of the correlation between subject’s pitch-naming ability and their principal wind instrument (the effect of playing a concert-pitch instrument versus a transposing instrument). Additionally, 1 year’s worth of training on a wind instrument represents a rather modest level of musical ability for children of this age; the use of more musically sophisticated subjects could have revealed much more. Nevertheless, a comparison of the results of the posttest of the training group with the generally established criteria of absolute pitch shows that, while no child’s identifications were completely errorless, a few seem to have displayed abilities comparable to native possessors of absolute pitch.

**Waters (1980)**

Joseph M. Waters, a student at the University of Minnesota, was intrigued by the apparent success of Brady (1970) in training himself to acquire absolute pitch. As we have seen, Brady not only proclaimed himself to be the first to do so, but claimed to have achieved this feat in only 2 months. Waters, a professional musician and nonpossessor of absolute pitch, sought to replicate several aspects of Brady’s training regimen as well and to add a few improvements of his own (Waters, 1980).
Waters analyzed Brady's training procedure and then reconstructed it, introducing new concepts one at a time. Seeking to develop a training method that could be used by subjects who had no previous knowledge of relative pitch, musical intervals, or music notation, he restricted the training pitch range to only two octaves, C₃ to B₄ (131 Hz to 494 Hz), instead of using Brady's three-octave range, A♯₂ to A₅ (117 Hz to 880 Hz). He divided training into 11 sections, one for each chroma other than C. Then he divided each section into six sequences.

The first sequence in each section involved only three tones, one of which was C₄ (262 Hz); the other two were nearby occurrences of the specific chroma being studied. There were, in effect, three different tones but only two chromata at this introductory level. These three tones were presented in a quasi-random pattern; the probability of the occurrence of C₄ was weighted at 50%. The second sequence and the ones which follow were only presented after the third section was reached. The second sequence in each section contained chromata from the preceding three sections placed in the fourth octave. C₄ (262 Hz) was always the lowest of the five tones and again it occurred 50% of the time. The third sequence in each section contained the same chromata as in the second sequence; yet, here the non-C₄ tones were placed in the third octave. C₄ was now the highest of the five tones and again it occurred 50% of the time. The fourth sequence in each section combined the spacings of the second and third sequences. It included 10 tones of five chromata and still maintained the same probability of C. In the fifth sequence, the chromata from all previous sections were added with C still weighted as before, while the sixth sequence presented them and C with equal probabilities of occurrence.
There was no prohibition regarding repeated notes during training sequences, and in early sequences repeated tones occurred quite often. Waters gave no indication about the order in which the chromata were introduced. He used sinusoids for training, their appearance randomized by computer, and recorded them on reel-to-reel tape. The tones were sounded for 1 second and the interstimulus time interval was held constant at 4 seconds. (Brady's interstimulus time interval had diminished from 5 seconds to 1.4 seconds as training progressed.)

To determine their absolute pitch ability, subjects were given a pretest developed by W. D. Ward, Professor of Communication Disorders at the University of Minnesota. The test consisted of 100 randomly selected tones presented in five blocks of 20 items. The test utilized only 10 of the 12 possible chromata. Waters noted that the pitch classes A and D were omitted "to aid statistical analysis of the results." (p. 8) Each of the 10 chromata appeared twice in every block but no chroma was allowed to follow itself immediately. With a diagram of the piano keyboard at hand, subjects were presented the 10 test tones, and told a number that identified each pitch. They were allowed to indicate these numbers on the keyboard diagram as a memory aid. The subjects' task was to write each tone's number on an answer sheet. Waters reported that the average score on this test by nonpossessors was approximately 1.4 bits of information transmitted, whereas perfect identification yielded 3.3 bits of information transmitted. Waters did not specify the registers, durations, or interstimulus time intervals of the sinusoidal test stimuli.

As for the training itself, subjects were provided a set of training tapes, and a book containing lists of the tones on the tapes and instructions
regarding the training techniques. Also provided were sheets to record the date, length, and comments regarding each training session. Subjects were instructed to remember the pitch C₄ and were permitted, at least in the beginning, to use their voice as an aid in its retention. (Brady had completely prohibited this technique in his training.) During testing, of course, subjects were not allowed to sing. Additionally, subjects were permitted to respond during training on an instrument, as long as they still attempted to recall C₄. Subjects were directed to practice daily, repeating each section for about a week, or until it was mastered. After 12 weeks of training, the subjects were posttested.

At the time of Waters' report (June, 1980) 3 subjects, himself included, were involved in his training experiment. All were either 26 or 27 years of age and preparing themselves for professional careers in music. He attributed the small number of subjects to a lack of available funding for subject participation and to the long duration of the training period. Apparently, several musically naive subjects began the experiment, reported initial success in following the training method, but dropped out due to lack of interest. This attrition was unfortunate considering the special measures Waters took to allow for their participation. Of the 3 subjects mentioned earlier, only 2 completed training and were subsequently posttested. After completion of all 12 sections of the training in the prescribed 12 weeks, one subject was able to correctly identify 85% of tones on the posttest. Waters described this performance only as representing "a good deal of improvement over the pre-test" (p. 18), but provided no pretest score for comparison. The sole remaining subject was Waters himself, and as his training differed
significantly from that of the other subject, the remainder of his report was largely anecdotal.

Waters added another form of training to the one utilizing quasi-random orderings of isolated pitches. This training was designed to improve one's "ability to name pitches within the context of tonal music." (p. 10) Participation in Waters' "Tonality Training" was optional, and was designed to be undertaken either simultaneously or after completion of the training with random pitches. At the time of his report, however, none of the other subjects had. As with the random-pitch training, "Tonality Training" contained 11 sections, one for each chroma other than C. Additional chromata were introduced in the same, still undisclosed, order as before. Waters first randomized the chromata to be used in a particular sequence. The chroma C was weighted by an unspecified amount, and these sequences were realized on a piano. For each chroma, a drill containing six of the seven steps of its major scale was played. The drill (expressed here in solfège syllables) was as follows: do, re, mi, fa, sol, fa, mi, re, do, ti, do. The submediant scale step la was omitted. The task was to identify the new chroma introduced on each trial by comparing it to the chroma C that they attempted to continuously imagine. The length of these sequences was not specified. The time interval between sequences was 5 seconds.

Prior to embarking on his training, Waters took the absolute pitch pretest in May of 1979. The test was administered by W. D. Ward, and Waters achieved an information transfer of 1.9 bits. Regarding the actual commencement of his training regimen Waters stated:

In my case, it is difficult to say exactly when my training began, since it took several months for me to settle on the final experimental design,
during which period I tried out various methods, at first using piano tones. (pp. 12-13)

Waters produced the final versions of both the sinusoidal random-note training tapes, and the piano-tone "Tonality Training" tapes, in September of 1979. He trained "consistently" (p. 13) from that time until his posttest in May of 1980—with the exception of 3 weeks in October of 1979. In the beginning of his training, Waters practiced primarily with the random-note portion of the experiment, and only occasionally with the "Tonality Training" portion. By late December of 1979, he had completed the random-note training and could identify non-weighted presentations of all 12 chromata in both octaves with only occasional semitone errors. He could not, however, recall the pitch of C; it had to be recognized from a stimulus such as the piano or a C-weighted training sequence. Nevertheless, Waters discontinued this portion of the training and afterwards focused entirely on "Tonality Training." By February of 1980 he could identify random presentations of all 12 "Tonalities" with only occasional semitone errors, and by March of 1980 he could recall C at will, without the aid of a stimulus.

In May of 1980, 12 months after the pretest, Waters was again tested by W. D. Ward. This posttest, however, included more tasks than the pretest. The first portion of the posttest contained 100 examples of the piano drill from the "Tonality Training" portion of the experiment. The examples had been taped by Waters himself 1 month prior to the test date and "were randomized according to the same stipulations employed in the design of the sine tone pretest."41 (p. 15) Ward started the tape at some arbitrary point near

---

41It is assumed that the 100 examples of piano drill were presented in five blocks of 20 items, and that no chroma could repeat itself. It is unclear whether or not the test utilized only 10 of the 12 possible chromata. The omission of two pitch classes (perhaps A and D again) is suggested by the total
its middle. Waters correctly identified the first 2 items, erred on the next 16 items by as much as a whole step in either direction, and then (after a presentation of the drill in E Major) correctly identified all the items on the remainder of the tape. Ward then rewound the tape to its beginning, and allowed Waters not only to respond to the previously unheard items at its beginning, but also to respond to the tape in its entirety. This time Waters correctly identified every item. Waters was then given the second part of the posttest, which was identical to the pretest he had taken a year earlier. He wrote:

I was slightly apprehensive about taking the test, since I remembered the trouble I had had with it previously. But after taking the first portion of the test, the chroma C was firmly implanted in my memory, and to my surprise, on the second time around the sine tone test proved to be extremely easy; I achieved a perfect score. (p. 16)

Two additional tests, with which Waters had no familiarity, were administered exactly 1 week later. The first test comprised random sine tones, the range, duration, and interstimulus time interval of which were not specified. Again, Waters achieved a perfect score. The second test involved 100 stimuli separated by quarter-tones. Fifty-five of Waters’ responses were correct, 44 responses were errors of a quarter-tone and the remaining error was of a semitone.

Waters reported that, “on all tests my answers were determined by making a relative judgment between the presented tone and the chroma C which I auralized [imagined] continuously during the test.” (p. 16) Nevertheless, he also noted that “there are certain chromata which seem to

number of test items, 100. Therefore, each of the 10 chromata probably appeared twice in each block.
stand out clearly, and which seem to have actually acquired an identity. At this time these chromata are $F\#$, $G$ and $A\#$.” (p. 17)

Waters was aware of the lack of scientific rigor that plagued this experiment. The extremely small number of subjects, use of self-experimentation, exclusion of certain chromata, lack of detailed practice records, equivocal testing procedures, and lack of a control group are just a few of the numerous flaws detailed above. Yet, Waters arrived at three conclusions. First, he asserted that his ability to name pitches out of context and to name the tonal center of short major-mode examples greatly improved following training. Second, he noted that it took him much longer to acquire absolute memory for pitch classes than the 2 months reported by Brady. And, lastly, he felt that his ability had not yet reached a plateau, and, in fact, would continue to improve.

Grebel'nik (1984)

In 1984, Russian researcher Stanislav G. Grebel'nik published several years’ worth of results from investigation of absolute pitch. Grebel'nik had undertaken seven preliminary experiments on the formation of absolute pitch in 34 children, ranging in age from 3 to 7 years, and had conducted a thorough review of the Russian literature in the field prior to publishing the results of his major training experiment (Grebel'nik, 1984). Grebel'nik held that while reaction to changes in pitch level was a natural, innate ability which could be measured immediately after birth, the ability to distinguish pitch levels (as in musical perception) first occur around age 3. Furthermore, he asserted that because the experience of musical perception is needed for the formation of absolute pitch, the formation of relative pitch must precede that
of absolute pitch. He also believed that active and passive absolute pitch do not always occur simultaneously, and that passive absolute pitch always forms before active. Grebel'nik proposed that absolute pitch could be formed in the course of perception of music organized as material for involuntary recall.

Based on his review of the Russian literature on absolute pitch, Grebel'nik formulated these five criteria for the determination of genuine absolute pitch. The first criterion is the ability to correctly identify randomly presented tones with greater than 63% accuracy. The second criterion is the absence of errors caused by the use of relative pitch. Third is the ability to identify pitch-classes in several octaves. Fourth is the identification of stimuli based on each tone's inherent individual qualities (chroma), and the last is the development of this ability in the course of ordinary musical activity.

Grebel'nik's major training experiment utilized 9 children ranging in age from 3 years and 9 months to 4 years and 2 months. The subjects were all students who attended kindergarten in the Fruzensk region of Leningrad from October of 1981 to May of 1982. Because the results of Grebel'nik's earlier experiments had suggested that absolute pitch formation was dependent on the subjects' prior level of relative pitch formation, he divided the subjects into two groups according to their levels of relative pitch ability. The 5 members of Group A were deemed to have a strong sense of pitch, and the 4 members of Group B were deemed to have a weak sense of pitch. No mention was made as to how these determinations were made. Both groups trained for two sessions per week, each session lasting 30 minutes. A total of
64 training sessions were conducted; 48 aimed at the formation of absolute pitch and 16 fostered its use in ordinary musical activity.

Before the training began, the subjects were given a pretest to measure their absolute pitch ability. Grebel'nik stated that after only two introductory sessions the children were able to learn the names of the tones produced on the white keys of the piano from C₄ to C₅. In a third introductory session, 20 tones from this collection were presented in random order for the children to identify aurally. To deter their use of relative pitch, Grebel'nik distracted the children by asking an extraneous question between presentations, which provided an interstimulus time interval of approximately 5 to 10 seconds. Percentages of correct responses ranged from 5% to 25% in Group A and 0% to 30% in Group B. Using the previously mentioned criteria it was quickly ascertained that none of the children possessed absolute pitch.

Grebel'nik used both Russian and Ukrainian folksongs as training material for his subjects. Adhering to a Russian theory of tonality, he assigned key qualities in relation to C Major, in this ascending order: C Major, D♭ Major, D Major, E♭ Major, E Minor, F Minor, G♭ Major, G Major, A♭ Major, A Minor, B♭ Minor and B Minor. Next, he found folksongs in these keys to serve as the basis for the training. For example, the Russian folksong *Poldiom, kum moi, vo lesochek* [Let's go, my friend, to the woods] is in the key of C Major, its melody begins and ends on C, and it was used to emphasize the *do* standard. Likewise, the Russian folksong *Ai, da golovushka moia bednaia* [Oh, poor me] is written with the d minor key signature, utilizes the hard dorian sixth (b-natural), its melody begins and ends on d, and it was used in training to emphasize the *re* standard. Each of the 12 songs was presented over the course of four sessions, and the children
trained together as a group. In the first session the children were told the pitches of the song, the meaning of the song’s text was explained, and pictures illustrating it were shown. The song’s melody was played on the piano in the original octave, then in different octaves following discussion. The second session reviewed the material of the first session, and a detailed analysis of the melody was given at the piano. The children were taught its tonal aspects (specific strengths and weakness of tonal implications were discussed) as well as its metric and rhythmic properties. The melodic structure as a whole was analyzed, and by the end of the second session the children could represent the song’s melodic line utilizing modeling hand motions. The third session reviewed the material of the second session, and focused on the harmonization of the song. This attribute was called the “musical color” of the song, and an association was established between the “color” and the “emotional tone of the tonality.” (p. 95) The general character of the piece was discussed along with its expressive character, and a detailed explanation of the harmonic decoration of the melody ensued. The fourth session reviewed the material of the third session, and then focused on a detailed harmonic analysis of the song. This “led the children to an understanding of tonal gravitation, the centrality of the tonic harmony, and the distinction of the main tone of a tonality.” (p. 95) Harmonized melodies were presented in various octaves to illustrate these points.

To help the children identify isolated pitches, Grebel’nik devised a series of 15 prompts, arranged in order of increasing difficulty, for each of the different songs associated with the 12 pitch standards. These reminders were designed to help the children extract the chromata of the pitches to be identified. After mastery of the first two songs (Session Nine), identification
trials were introduced during training. A tone would be sounded, and the child was directed to identify it. If the child struggled to answer, (s)he was given prompts until correctly identifying the song associated with the standard pitch. Gradually the names of the songs were replaced by solfège syllables.

Following the 48 training sessions dedicated to the formation of absolute pitch, the subjects were given a posttest that was considerably harder than the pretest. Each child had to identify the pitch names of 20 randomly ordered piano tones presented from among the chromatic—not just the diatonic—pitches, ranging from B₂ to F₆. The children were not required to identify the octave, and no prompting or correction occurred. The percentage of correct responses in Group A (the strong group) ranged from 60% to 100%: those in Group B (the weak group) ranged from 20% to 50%. The mean percentage of correct responses for was 78% for Group A; 34% for Group B. Individual test scores for each subject are shown in Table 10.

Table 10. The percentage of correct pretest and posttest responses of the nine preschool children trained by Grebel'nik (1984).

<table>
<thead>
<tr>
<th>Group</th>
<th>Name</th>
<th>Gender</th>
<th>Pretest</th>
<th>Posttest</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Olya N.</td>
<td>Female</td>
<td>5%</td>
<td>60%</td>
</tr>
<tr>
<td>A</td>
<td>Masha E.</td>
<td>Female</td>
<td>15%</td>
<td>100%</td>
</tr>
<tr>
<td>A</td>
<td>Tanya M.</td>
<td>Female</td>
<td>25%</td>
<td>80%</td>
</tr>
<tr>
<td>A</td>
<td>Alyosha G.</td>
<td>Male</td>
<td>15%</td>
<td>75%</td>
</tr>
<tr>
<td>A</td>
<td>Anfisa L.</td>
<td>Female</td>
<td>20%</td>
<td>75%</td>
</tr>
<tr>
<td>B</td>
<td>Olya S.</td>
<td>Female</td>
<td>30%</td>
<td>40%</td>
</tr>
<tr>
<td>B</td>
<td>Anya T.</td>
<td>Female</td>
<td>10%</td>
<td>20%</td>
</tr>
<tr>
<td>B</td>
<td>Roma M.</td>
<td>Male</td>
<td>0%</td>
<td>25%</td>
</tr>
<tr>
<td>B</td>
<td>Alyosha B.</td>
<td>Male</td>
<td>20%</td>
<td>50%</td>
</tr>
</tbody>
</table>

The children next undertook a series of training sessions aimed at applying their newly acquired skills in a musical context. They practiced
singing, in the correct keys, the exercises of a Russian solfège text. They started in C Major, and later moved to other keys. After 16 sessions the children were again administered the posttest described above. The results were about the same. The percentage of correct responses in Group A ranged from 65% to 100%; those in Group B ranged from 20% to 50%. The mean percentage of correct responses for was 85% for Group A; 36% for Group B. Individual scores were reported for some, but not all, of the subjects.

Grebel'nik concluded that genuine absolute pitch was developed in the strong group according to all five of the previously described criteria. Progress was reported for the weak group as well, but clearly the standard was lower. Grebel'nik also concluded that the formation of relative pitch precedes that of absolute pitch and that musical activity not only supported, but perfected the formation of the latter.

Grebel'nik’s study represents a significant advance in the study of the formation of absolute pitch in young children. His apparent success is quite impressive and his findings are difficult to dispute. It would be worthwhile to replicate this study, possibly improving upon the quality of Grebel'nik’s work by adding a control group, by training a larger number of subjects, and by using longer, more thorough test instruments. Unfortunately, this study of young children sheds no light on the subject of training absolute pitch in older subjects.

**Faivre (1986)**

At the time of this study, the most recent investigation of the acquisition of absolute pitch was undertaken by Irene A. Faivre at the
University of Colorado at Boulder.\textsuperscript{42} Faivre (1986), recognized that judgments of a tone's chroma are much subtler than those of its height, and noted that while most musicians consider tones an octave apart as equivalent, most musically naive subjects fail to recognize chroma spontaneously. Faivre wrote:

In terms of Van Noorden's (1982) physiological model, the problem for the naive subject in the AJ [Absolute Judgment] task is that he or she attends exclusively to the information from the channel which admits pitch height information and ignores the channel which admits chroma information. Most training paradigms encourage the subject to attend even more closely to pitch height and do nothing to develop awareness of pitch chroma. A more effective training strategy would emphasize information from the chroma channel. Subjects who recognize chroma information could build a more efficient mental representation of the stimuli, one based on two dimensions rather than one. This was the basic assumption of this present study. (Faivre, 1986, p. 17)

Faivre's training experiment utilized 4 subjects;\textsuperscript{43} 2 were musicians, and 2 had little or no musical background. One of the musicians, Subject MB,

\textsuperscript{42}Faivre first wrote on the topic of absolute pitch while she was a master's student in psychology. Her thesis experiment (Faivre, 1983) tested 12 subjects ability to verbally identify seven numbered sinusoidal tones. The subjects were students from an introductory psychology class; none were musicians. The frequencies of the sinusoids corresponded to the pitches from C\textsuperscript{#4} to G\textsuperscript{4}, inclusive. They participated in a 224-trial training session, a control condition, and two 72-trial experimental conditions. Feedback was provided, and various types of interference tones were included during different experimental conditions. The experiment dealt primarily with disrupting the subjects ability to make comparisons with previous stimuli held in their short-term memories, rather than the effect of their training on the acquisition of absolute pitch.

\textsuperscript{43}A fifth subject, a female freshman whose major and musical ability were not specified, also began the training. Faivre reported that she was dropped from the experiment after it was determined that the subject "was unable to reliably identify four musical stimuli in an absolute judgment task...This subject's data were not analyzed." (Faivre, 1986, p. 23) Only test data for the other four subjects were reported.
was a female violinist in her senior year of college. All of the members of her family were musicians, and her brother and sister possessed absolute pitch. She was reported to have very good relative pitch, and already possessed an absolute memory for A4 prior to the beginning of the experiment. This ability was verified in a frequency setting task using a computer-generated tone. Her skill was assumed to be the result of practice in tuning her violin to “concert A” since the age of two. The other musician, Subject MC, was a male doctoral student and Teaching Assistant in the music department. His primary area of expertise was the piano, and he was skilled on other keyboard instruments as well. As for the two nonmusicians, Subject NP was a senior from the psychology department. He had never received formal musical training. The other nonmusical subject, Subject NV, was a female graduate student, also from the psychology department, who specialized in psychophysics. As a child she had received 2 years of piano lessons, but had played only rarely in recent years.

The training and testing stimuli used in the experiment consisted of 12 different tones. However, instead of encompassing 12 unique chromata, they encompassed only four. Faivre’s stimuli comprised a fully-diminished seventh chord in three octaves. The pitches used were C3, D#3, F#3, A3, C4, D#4, F#4, A4, C5, D#5, F#5, and A5. The fundamental frequencies of these pitches range from 131 Hz to 880 Hz. The tones themselves were triangle waves generated on a Commodore 64 personal computer. Tones were amplified and presented binaurally to subjects through headphones at equal loudness levels. The overall loudness was controlled by each subject.

Faivre’s training regimen included nine phases. These were arranged into three sections of three units each. Sections were distinguished by the
addition of new chromata. The first section contained two chromata (D# and F#), the second section contained three chromata (D#, F#, and A), and the third section contained all four chromata (D#, F#, A and C). Units were distinguished by the addition of a new register. The first unit in each section utilized the fourth octave (USA Std.), the second unit utilized the fourth and fifth octaves, and the third unit utilized the third, fourth and fifth octaves. Subjects advanced to the next phase of training when they could maintain an accuracy level greater than 90% correct of the stimuli presented during a session, for three consecutive sessions. The stimuli presented at each phase of training are shown in Table 11.

Table 11. The order and arrangement of the 12 stimuli presented at each of the nine training phases in Faiivre (1986).

<table>
<thead>
<tr>
<th>Phase</th>
<th>Third Octave</th>
<th>Fourth Octave</th>
<th>Fifth Octave</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>D#4 F#4</td>
<td></td>
<td>D#5 F#5</td>
</tr>
<tr>
<td>1.2</td>
<td>D#4 F#4</td>
<td>D#5 F#5</td>
<td></td>
</tr>
<tr>
<td>1.3</td>
<td>D#3 F#3</td>
<td>D#4 F#4</td>
<td>D#5 F#5</td>
</tr>
<tr>
<td>2.1</td>
<td>D#4 F#4</td>
<td>A4</td>
<td>D#5 F#5 A5</td>
</tr>
<tr>
<td>2.2</td>
<td>D#4 F#4</td>
<td>A4</td>
<td>D#5 F#5 A5</td>
</tr>
<tr>
<td>2.3</td>
<td>D#3 F#3 A3</td>
<td>D#4 F#4 A4</td>
<td>D#5 F#5 A5</td>
</tr>
<tr>
<td>3.1</td>
<td>C4 D#4 F#4 A4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.2</td>
<td>C4 D#4 F#4 A4</td>
<td>C5 D#5 F#5 A5</td>
<td></td>
</tr>
<tr>
<td>3.3</td>
<td>C3 D#3 F#3 A3</td>
<td>C4 D#4 F#4 A4</td>
<td>C5 D#5 F#5 A5</td>
</tr>
</tbody>
</table>

Faiivre wrote that training sessions lasted approximately 1/2 an hour, and usually two consecutive sessions were run each day. Subjects trained individually using tones presented by the computer. The pace of the session
was determined by the length of time it took for the subject to respond to the stimuli, and this time interval was unlimited. The duration of each stimulus tone was 1 second at the beginning of the training, but was reduced to 0.5 seconds as the subjects gained proficiency. Each session began with a familiarization period before commencement of training. The actual training sequence began with an audible click followed 1 second later by a visual cue to notify the subject. Following 2 seconds of silence, the stimulus tone was presented. Immediately after a response, a 2-second feedback period occurred in which subjects were signaled they were correct, or given the answer. A 5-second interstimulus time interval followed before the next presentation. The response task consisted of pressing one of four keys on the computer keyboard to identify the pitch class, and pressing one of three keys to identify the octave. Appropriate labels were attached to the keyboard only for the particular responses required by each phase of the training. The entire training regimen was complete when the subject could maintain an accuracy level of 90% correct on Phase 3.3 for three consecutive training sessions, a total of 216 trials. Trials were grouped into blocks, each containing two randomized presentations of each stimuli contained in the phase being practiced. The majority of sessions contained at least three blocks.

To measure the effect of training, test sessions were scheduled after each of the three sections of training. During tests, subjects identified the stimuli by pressing the numbers 1 through 12 on the keyboard. No feedback was provided, and five interference tones similar to the stimulus tones were played between judgments. The fundamental frequencies of the computer-

44Specifically, the pitch classes C, D#, F#, and A were indicated on the keyboard by the 2, 3, 4, and 5 keys, respectively. Similarly, the third, fourth, and fifth octaves were indicated by typing 6, 7 and 8, respectively.
generated interference tones were determined by a random number generating program, and ranged from 66 Hz to 1760 Hz (one octave above and below the stimulus range). In reality, testing occurred only following the second and third sections.

Three of the 4 subjects who participated in the study were able to obtain an accuracy level of 90% correct for four chromata in three octaves. The musicians, Subjects MB and MC, were able to meet the criterion in 24—the minimum possible—and 29 sessions, respectively. As for the nonmusicians, Subject NP was able to meet the criterion in 65 sessions, but Subject NV was still unable, in 56 sessions, to advance to the second section—that is, she could not yet reliably distinguish the chromata D♯ and F♯ in three octaves. Faivre wrote that “data from the two musicians were inconclusive, partly because of insufficient data and partly because they appeared to have additional strategies available to them which were based directly on their musical training.” (Faivre, 1986, p. 137) Having only one remaining subject who completed the training, Faivre wrote “it is apparent that the data of NP provide the most information about learning in the AJ task. Therefore the formal tests of the models were conducted using his data.” (Faivre, 1986, p. 37) The rest of the report details Subject NP’s progress and performance throughout training (perhaps adding up to no more than the acquisition of rudimentary relative pitch skills) and compares them to a proposed two-dimensional “Matrix Model” of the acquisition of skill in the absolute judgment of pitch.

Since Faivre’s stimuli represent a symmetrical collection, the objective of her training experiment (the absolute identification of four chromata in three octaves) was somewhat more ambitious than that of Cuddy’s 1971
experimentation. Nonetheless, the ability displayed by Faivre’s subjects at the conclusion of their training regimen was very modest indeed when compared with the abilities of actual possessors of absolute pitch. In fact, it is not certain that Faivre’s subjects could absolutely identify the four chromata in question, because they were never tested in a context containing any of the other eight. This limitation, when combined with a small number of subjects, lack of a control group, lack of a pretest, the allowance of a subject to participate who already possessed some of the skills about to be learned, and general disregard of the data derived from actual musicians leaves almost no basis for meaningful conclusions.

Summary

Historically, absolute pitch has been one of the main topics in the field of the psychology of music; so much so, that over the course of the last 18 months, well over 300 sources have been catalogued in the present study. Included among them are the 30 reports of previous attempts to develop absolute pitch reviewed here. While experimenters have employed a myriad of training techniques, several central themes run throughout.

One popular technique has emphasized the acquisition of a single internal pitch standard; this technique is founded on the notion that if absolute pitch ability could first be fostered for a single pitch, other tones could later be acquired. Salomon Jadassohn was the first, in 1899, to suggest this single-standard technique. Jadassohn advocated “acquiring the possession of a fundamental tone,” from which all other tones may be derived, and recommended the use of a tuning fork to help fix it in mind. Fifty years later, Paul Hindemith made these same suggestions in his
Elementary Training for Musicians (1949). Another composer, Julien Falk, claimed in 1975 that absolute pitch could be acquired through the practice of solfège and dictation exercises. Unfortunately, their accounts are purely anecdotal, offering no experimental evidence to back their claims of success.

In 1916, Evelyn Fletcher Copp stated that children who had been considered devoid of musical talent have, after a few months of training, been able to sing on demand middle C, to recognize it when it is played or sung, and soon thereafter to add the other musical tones with the same results. She acknowledged that the rate of progress varies, and that not every child can acquire absolute pitch by training, but she asserted that most can do so if they begin at an early age. Copp wrote, "by proper training this power may be acquired, speaking very conservatively, by 80% of normal children."

In 1965, Fred Fisher suggested that students wishing to acquire absolute pitch begin by gradually lengthening the recall time for a single tone. In 1925, Helen Mull had gathered empirical evidence for this position. In one of her three experiments, subjects were presented 15-minute doses of middle C in hour-long, weekly training sessions for most of the academic year. She reported an average increase from 52% to 89% of correct responses to C diatonic test stimuli following training.

A variation of this technique advocated the learning of a set of multiple reference standards, specifically the tones of a single triad, with hopes that other tones or triads could be mastered later. This design was first proposed and tested in 1916 by a student of C. A. Ruckmick at the University of Illinois using, naturally, a C Major triad. Training experiments based on mastery of a triad were still being tested by Lola Cuddy as recently as 1971. However, the most remarkable triad-training experiment was conducted in
Japan in 1940. Twice each week Aiko Komatsu played triads on the piano to a 5-year-old girl and after each presentation she read, in German, the names of the tones, from lowest to highest for the subject to repeat. After about 15 months of training the subject could name the pitch of the root and the quality of a chord immediately upon hearing it, and could also notate her answer. During training the subject’s percentage of correct responses varied from a low of 64%, to a high of 96%. Interestingly, the child always responded in a low-pitched voice to low-register piano chords, and in a high-pitched voice to ones of the upper registers.

Instead of training subjects to recall specific triads to aid in absolute pitch judgments, recent experimentation has dealt with the recall of specific musical compositions. In 1984, Soviet researcher Stanislav Grebel’nik trained nine kindergartners, using Russian and Ukrainian folksongs in specific keys as their training material. After 32 weeks of training sessions, one group’s mean percentage of correct responses to random piano tones rose from 16% to 85%.

The idea that unguided practice on numerous tones would prove helpful in acquiring absolute pitch is in direct opposition to techniques emphasizing “standards-training.” In 1895, Max Meyer and colleague Victor Heyfelder trained themselves, gradually adding more and more tones, without placing any emphasis on a central tone or particular standard. Evelyn Gough, in 1922, had subjects practice naming piano tones in four 10-minute sessions per week for most of a school-year. Her subjects were instructed to devise their own training techniques, freely and independently creating them as they progressed.
The advent of improved technology led in the 1960s to a resurgence of this "random-training" technique. Experiments devised by Robert Lundin and Joseph Allen, and later by Michael Terman, employed teaching machines similar to those used in programmed learning. These machines allowed their subjects to train individually and to receive correction immediately. Michael Vianello critically evaluated the role of feedback. In 1978, Lisbeth Hurni-Schlegel and Alfred Lang used a specially constructed training device to present randomly selected square wave tones at the frequencies of the 12 equal-tempered semitones from C4 to B4. Their results, using 21 children, indicate that absolute pitch identification can be improved through simple test repetition within the short training period of 5 days, with only 15 minutes of practice daily.

A third class of psychological experimentation in absolute pitch training is more concerned with absolute judgments of pitch, than with what musicians would normally call absolute pitch behavior. By this it is meant that as basic research, neither practical application, nor normal usage is of immediate interest, and the constraints of using the stimuli of common musical usage are waived. The 1934 work of Carl Wedell is a good example of such basic psychological experimentation. Wedell constructed various divisions of a frequency range according to approximately equal numbers of difference limens, not musical intervals. He concluded that the number of frequencies a subject could be trained to identify fell somewhere between 9 and 13. Irwin Pollack, the first investigator to apply the theory of information to the absolute judgment of pitch, concluded in 1952 that response accuracy in the absolute judgment of tones in a series is relatively invariant, regardless of the number of tones, their frequency range, or their distribution. Two years
later, E. B. Hartman, integrating elements of these previous experiments, concluded that in the absolute judgment of tones, pitch spacing was positively related to the number of tones that could be learned, rapidity of learning, and long-term retention. As recently as 1970, Cuddy determined the effect of certain variables, such as subjects' musical ability and the type of training technique employed, using sets of nine sinusoidal stimuli spaced according to their pitch distance in mels.

A simpler, pragmatic "mental" approach was advocated by several musicians. In 1925, Soviet researcher E. A. Mal'tseva trained subjects to identify the names of 36 tones from all 12 pitch classes in the three middle octaves of the piano. She instructed subjects to form mental images of the tones they were about to judge, and to focus their attention systematically on the timbre rather than on the pitch height of piano tones. In 1954, Maury Deutsch thought that absolute pitch was a mental process independent of other faculties. He instructed students to visualize a single tone so firmly that it could be recalled at will. After the successful visualization of the tone, the student was instructed to concentrate on its letter name, in the hope that, eventually, the thought of just the letter name would evoke the tone's visualization. Deutsch himself likened this first step to Pavlov's famous conditioning studies. Unlike Mal'tseva, Deutsch provided no experimental evidence regarding the success of his method.

The most remarkable training technique dependent upon one's mental powers was the chromesthetic training for absolute pitch advocated in 1919 by Edward Maryon. Maryon developed and codified his theories correlating color and sound in the teaching of music, naming his technique Marcotone. He held that because the source of both light and sound is vibration, the two
co-exist in a natural relationship. Believing in the law of atavism, he taught that through the process of color-thought, specific ratios of visual color can be correlated to specific ratios of tone. In other words, if a color is mentally retained, the brain will signal this fact with a response of the exact corresponding tonal frequency from the mouth.

The most recent category of training experiments have explored a weighted training technique. This technique, first attempted in 1968 by Cuddy, expands upon the learning of a single standard by its repeated presentation in increasingly difficult settings. Cuddy’s subjects practiced with four levels of training tapes, each containing decreasing portions of the standard. Cuddy reported that all subjects improved after training and that 2 subjects achieved perfect scores on tests of the A4-standard. A follow-up experiment found this training method was more effective than ordinary practice with correction. In 1970, Paul Brady implemented techniques adapted from Cuddy in a remarkable self-training experiment. After training for 1/2 an hour daily for 2 months, he gradually began to retain the unique chroma of each pitch class. On a posttest, Brady correctly identified over 96% of the test tones within plus or minus one semitone. While a 1972 attempt to replicate Cuddy’s findings by Heller & Auerbach found the weighted training method to be no more effective than ordinary practice with correction, the success of Brady’s C-weighted training was replicated by Waters in 1980, though Waters noted his progress was much slower.

Surprisingly, only two training experiments have focused on the recognition of tone chroma. The first such experiment was conducted in 1915 by Wolfgang Köhler. He asserted that absolute pitch was not based on memory for tone height, but the memory of other tonal characteristics which
he described as the tone body. He believed that the two were independent and, after only 2 weeks of practice focusing exclusively on the "body" of the tone, could correctly identify 51% of piano tones from a five-octave C diatonic collection. A recent CAI experiment by Irene Faivre reports success with this chroma technique, especially with musicians, but her experimental goal, mastery of 4 of the 12 chromata, was modest indeed when compared with the abilities of actual possessors of absolute pitch.
CHAPTER III

PROBLEM

The Role of Chroma in Absolute Pitch

Only during the last 150 years have experimenters in music psychology formally recognized the perceptual similarity of tones separated by octaves. This perceptual similarity is usually referred to as octave generalization. Of course, musicians have been aware of this similarity for centuries, and have treated tones separated by octaves as being, for the most part, equivalent. Indeed, the repeating note names of Western music reflects this relationship. Numerous investigators have studied the human perception of octave generalization in humans, especially by musicians (Humphreys, 1939; Allen, 1967; Deutsch, 1972c, 1973b; House, 1977; Thurlow & Erchul, 1977; Kallman, 1982; Ohgushi, 1983; Deutsch, Moore, & Dolson, 1986; Demany & Semal, 1988). Some have offered evidence to suggest that octave generalization exists in infants (Demany & Armand, 1984) and even animals (Blackwell & Schlossberg, 1943). The predominant opinion, however, is that octave generalization is probably learned (Burns & Ward, 1982). Moreover, several writers have noted that persons with absolute pitch sometimes correctly identify the pitch class name of a tone, but place it in the wrong octave (Stumpf, 1883; Baird, 1917; Bachem, 1937, 1954; Révész, 1954). The common notion that links both issues is that pitch can best be described as a

124
combination of two separate attributes, tone height and tone chroma. Tone height is related directly to the frequency of the tone, and can be thought of as the overall highness or lowness of a tone in terms of pitch level; it is a linear function. Tone chroma is related to the position of the tone within the ordered series of pitch classes within the octave; it is a circular function. Tone chroma is the attribute of pitch whereby tones are perceived as more closely related to their octaves than to any other tones. This quality may be thought of as the characteristic C-ness of a C, or the C#-ness of a C#, etc., and was at one time referred to as tonality (Meyer, 1914; English & English, 1958). Deutsch (1980) acknowledges that this attribute, termed tone chroma by psychologists, is termed pitch class by music theorists. Thus, in the vocabulary of music theorists, information concerning tone height is one of the features that distinguishes pitch from pitch class.

The idea that there is more than one perceptual attribute for pitch is not a new one. Various writers have proposed helices or other representations of their structural conceptions of pitch (Droatisch, 1855; Ruckmick, 1929; Révész, 1913/1954). Figure 5 reproduces the first helical representation of the bidimensional attributes of pitch.
Figure 5. Digitally enhanced reproduction of Drobisch's helical representation of musical pitch (1855).
Shepard (1964, 1982) is generally credited with crystallizing much of the work in this area and, by means of his endlessly rising or falling “Shepard’s tones,” with devising a way in which the perceptual bidimensionality of tone height and tone chroma could be experimentally tested. In his three-dimensional, regular (non-distorted) pitch spiral, tones separated by an interval of an octave lie closest together within each turn of the helix. The associated terminology became standardized much later than the introduction of Drobisch’s helix, and the cyclical nature of tones came to be known as tone chroma, or simply chroma (Bachem, 1950; Révész, 1913/1954; Shepard, 1964, 1982; Risset, 1978).

The most recent explanations of absolute pitch behavior support the bidimensionality of pitch (Ward & Burns, 1982; Van Noorden, 1982). It is generally believed that persons having absolute pitch perceive pitches differently than do those lacking absolute pitch. Nonpossessors identify pitches largely, if not solely, on the basis of learned relations of pitch height—i.e., relative pitch. While absolute pitch possessors may use relative pitch skills, they have the additional ability to form mental representations of pitches based on standards of pitch chroma stored in long-term memory. When asked to judge chroma, absolute pitch possessors are far more accurate than nonpossessors. When judging the octave in which a pitch lies, absolute pitch possessors are much like nonpossessors (Carroll, 1975). The explanation of absolute pitch ability in terms of a two-dimensional model of pitch accounts for the frequently mentioned octave errors of absolute pitch possessors.
Selection of the Training Method

At the time of this study, only three tutorials for the acquisition of absolute pitch were commercially available. One was currently being distributed by Dr. Maury Deutsch, whose journal articles (Deutsch 1954, 1986) were discussed in Chapter 2. Deutsch distributed a catalog of original study materials primarily related to the areas of composition and arranging, and also made available several collections of articles and lecture notes. Among those in his Series B are (Deutsch, 1970), Elements of Solfeggio: Sight Singing & Absolute Pitch and (Deutsch, 1974), Study of Absolute Pitch. The latter claims to include "an intensive study of technique necessary for the eventual ability to recreate pitches mentally (absolute pitch)."¹ The study guide expands upon Deutsch's "visualization" technique, and includes a practice schedule requiring a minimum of one and a half years. In Series L, a supplementary article is entitled Simulated Absolute Pitch. Though each series is available for a modest fee, it was not found that the techniques are in widespread use.


¹The catalog is available from Dr. Deutsch, 150 West 87th Street, Suite 7C, New York, N. Y. 10024, Phone: (212) 724-4722.
range from very easy to impossibly difficult. A search of O.C.L.C. location records revealed that copies of Falk's treatise can be found in over a dozen research libraries but, as with Deutsch's method, it could not be determined that his techniques were in widespread use.

The third, and most recent, method of absolute pitch instruction found was by David L. Burge; it is entitled Perfect Pitch: Color Hearing for Expanded Musical Awareness (Presque Isle, MI: Innersphere Music Studio, 1983). A report by Burge's publisher (American Educational Music Publications, Inc., 1986a) advertises:

David L. Burge's Perfect Pitch Ear-Training Course is the most popular and widely listened-to music course today! . . . Burge has presented his revolutionary Perfect Pitch Seminar™ at distinguished schools in both the U.S. and Canada. (p. 1) . . . World famous for his Perfect Pitch Seminar, Burge explains how to gain perfect pitch. His simple teaching is for all musicians. (p. 2) . . . Burge's Color Hearing Technique™ takes only about 15 minutes of listening daily, and no previous musical experience is necessary! . . . It's easy! You just listen to the tapes and read Burge's short handbook. Your home-study Course is the exact same Course Burge has presented at colleges to musicians from rock guitarists to symphony violinists. (p. 3) . . . Burge is applauded around the world for his simple technique to develop perfect pitch. . . . Burge's tapes are heard in over 30 countries worldwide. (p. 4)

A search of O.C.L.C. location records revealed that copies of Burge's works can be found in nearly 150 research libraries. While not attempting to validate any of these claims of international notoriety, the use of Burge's method was assumed, prima facie, to be more widespread than either of the other methods, and was the one chosen for further investigation.

---

2O.C.L.C. is an acronym which stands for Online Computer Library Center, a company which indexes the holdings of over 9,400 large research libraries.
Burge’s Technique

Mr. Burge, a native of Wilmington, Delaware, describes himself as a “keyboard artist” and an “author/composer.” He attended classes in music theory at the Wilmington Music School, the University of Delaware, and the Wheaton Conservatory of Music in Illinois (Burge, 1984b). He makes no claims of graduation, and is not affiliated with any research facility or institution of higher learning. The techniques of his absolute pitch method for adults are encompassed in a handbook (Burge, 1983) and a series of cassette tapes (Burge, 1986a, 1986b, 1987). The handbook, *Perfect Pitch: Color Hearing for Expanded Musical Awareness*, is a short 60-page monograph. It is written in a non-technical, non-scholarly style and, in fact, its very title contains the imprecise and confusing terms—perfect pitch and color hearing. Burge’s use of the term color hearing is clearly inconsistent with the more frequently cited writings of Ostwald (1964) and Peacock (1985), and is especially

---

3It is appropriate to clarify an often confused matter of identity. This David Burge, a young man in his early 30s and author of an absolute pitch training method (Burge, 1983), is not Dr. David Burge, the Grammy Award-winning Chairman of the Eastman School of Music piano department (College Music Society, 1987; Press, 1985).

4 For example, lexicographers and physicists have long disdained the use of the term perfect pitch in favor of absolute pitch; the former is thought to evoke the notion that the faculty is infallible, while the latter more accurately denotes that the faculty merely involves judgments of an absolute, not relative, nature (Slonimsky, 1930; English & English, 1958; Hall, 1982). Even more significantly, the term color hearing (color-hearing, colored hearing, Farbenhören, audition colorée) has taken on a precise meaning in the field of music psychology as the most common type of chrom(a)esthesia or chrom(a)esthesia. Chromesthesia is the most common variety of synesthesia, and has long been the topic of scientific inquiry (Myers, 1911; Myers, 1914; Kelly, 1934; Howells, 1944; Carroll & Greenberg, 1961; Ostwald, 1964; Marks, 1975a; McCluskey, 1975; Critchley, 1977; Haack & Radocy, 1981; Polzella, Kuna, Biers, & DaPolito, 1982; Cuddy, 1985a; Peacock, 1985; Bernard, 1986). Color hearing is defined as the vivid and persistent association of specific visual colors with sounds, spoken words or letters, or any other auditory sensation.
misleading because earlier absolute pitch training methods (Maryon, 1919, 1924) actually were devoted to its acquisition through the process of color hearing. Burge rationalizes his incongruous terminology with this cryptic statement:

It may have been agreeable not to have added another definition of color to the musician's already crowded list, but the sense of perfect pitch is master of this word and therefore rightfully deserves its place. (p. 19)

To access the originality of his contributions, it was necessary to determine what exactly Burge means by his idiosyncratic use of the term "color hearing." He writes:

Color hearing is not to be confused with synesthesia or color association, nor is it in any way a reference to instrumental "tone color" or other musical definitions of the word color. Color hearing simply means the facility to "hear in color," which implies the ability to identify musical frequencies (pitches) by their perceived pitch color.

The color of a pitch could be described as a certain "quality" of sound which is peculiar to and characteristic of each musical tone. To most ears, there is apparently nothing different about one pitch from another except its "highness" or "lowness." To the color ear there is a subtle, yet distinct difference between each individual pitch—some quality which distinguishes one from the next. "Highness" or "lowness" does not enter into pitch identification—color alone is the secret of the ear's ability to know a pitch. (p. 28)

The different octaves of a pitch could be likened to different "tints" of that same color. Every F#, for example, has the characteristic "twangy" color of an F#, no matter which octave it occurs in. The different octaves will simply determine if this color of F# is "brighter" (higher) or "darker" (lower). (p. 39)

A careful rereading of these excerpts substituting the word "chroma" for the words "pitch color," and the words "the ability to identify chroma" for the words "color hearing" effects little or no change in their meanings while bringing them in line with the accepted terminology. Clearly, Burge is referring to the ability of possessors of absolute pitch to perceive tone
chroma, in addition to tone height. Burge's imprecision may merely be the result of a lack of familiarity with the related literature. Regarding the body of knowledge concerning absolute pitch, Burge (1984b) writes "very little had [sic] been written on the subject—mostly just old research experiments in the '30s and '40s which don't really conclude much." (p. 2) Likewise, Burge's review of previous attempts to train absolute pitch includes the work of only two writers (Mull, 1925; Hindemith, 1946), and the entire handbook cites only 14 different sources. Indeed, the cover notes proclaim it to be "the only ear-training technique handbook for developing absolute pitch!"

Burge's aural-training technique comprises several simple activities. These include: (a) practice with a partner, or less preferably by oneself (never with a tape recorder); (b) practice with the same timbre until the final stages of training, preferably using the piano or one's primary instrument; (c) practice with an in-tune instrument, although one need not become obsessive about it; (d) practice unaided by subvocalization, that is, the comparison of the relative changes of muscular tension in one's vocal cords could possibly aid pitch differentiation; (e) practice ordinarily limited to one session daily, usually 15 to 25 minutes in length; (f) practice with one's focus directed toward color, not relative pitch; and (g) practice in a relaxed atmosphere, one devoid of tension or fatigue. Burge's handbook contains a total of 18 mastery-based exercises which comprise 10 of the handbook's 60 pages. Advancement to the next exercise is allowed after achievement of a high degree of accuracy, usually 95% correct, or 20 consecutive correct identifications. Following completion of Exercise 14, students are expected to possess the ability to identify single tones, and two- and three-tone simultaneities randomly selected from all 12 chromata in the second, third,
fourth, and fifth octaves by their pitch class names. Study of the chromata is undertaken in the following order; E♭, F♯, A, C, C♯, E, G, B♭, D, F, A♭, and B. The final four exercises refine and expand one's new abilities by adding octave identifications to those of chroma, expanding the range of identifiable tones to that of the piano, and increasing the number of identifiable timbres.
Hypotheses

An experiment was devised to test whether college musicians could acquire absolute pitch through training. A commercially available training method (Burge, 1983) which appeared to be in widespread use, but which had never been scientifically tested, was selected. The absolute pitch-naming abilities of a group of volunteers, all nonpossessors of absolute pitch, were measured. The volunteers became the experimental group and trained for an academic year utilizing a method that focused on acquiring chroma judging skill. An equivalent control group was established to determine whether nonpossessors’ abilities change without instruction over the same time period.

Because the selected training method emphasized attention to chroma, it was deemed to have excellent chances for success. Overall, several predictions were made:

1. Nonpossessors of absolute pitch would lack chroma identification skills.
2. Their level of ability to identify tone chromata would not increase without training.
3. Subjects who trained would be superior at this skill to those who did not train.
4. Advancement in the training method would be related to the amount of effort expended by the subjects.
5. Chroma identification ability would be related to advancement in the training method.
CHAPTER IV

METHODOLOGY

Experimental Design

In an ideal situation, studies would meet the criteria of a "true experiment" as defined by Campbell and Stanley in *Experimental and Quasi-experimental Designs for Research* (1963). The best design for experiments such as this is either the posttest-only control group design, or the pretest-posttest control group design. The first design assumes that both the experimental group and the control group are equivalent with respect to the dependent variable, in this case, absolute pitch naming ability. Because such an expectation was unreasonable and the study required a long period of experimental manipulation, the use of the second design was mandated. The only adverse effect of including a pretest would be pretest-posttest interactions—threats to the external validity of the study caused by the

---

1In an ideal experimental setting, unlimited numbers of willing and equivalent subjects would experience only identical events between test measurements, the passage of time would have no bearing, the effects of taking a test would be negligible, all tests would be exactly replicated and uniformly scored, samples would be unbiased and totally indicative of their representative populations, no differential loss of subjects would occur, and no element of experimental design could confound the effects of the experimental variable. Unfortunately, such ideal conditions exist only in the minds of dreaming experimenters and, as a result, strict accordance to accepted experimental designs allows our only hope of obtaining adequate and proper data.
sensitizing effects of the test itself. In this study, such interactions would likely be minimal because of the nonreactive nature of an absolute pitch pretest and the long period between the tests.

Both “true experimental designs” call for the random selection of subjects from a sample population, the random assignment of subjects to two groups, and the random assignment of the treatment to one of the two groups. Obviously, a subject’s year-long participation could not be expected merely because his name had been randomly selected and assigned to the group that happened to be the experimental one—subject mortality would be prohibitive! Because experimentation in musical perception is rarely, if ever, suited for these types of randomization procedures, use of a “quasi-experimental design” was required.²

The preferred design for this study was the nonequivalent control group design, probably the most widespread experimental design used in educational research today. Even a design implementing a nonequivalent control group is considered far superior to its alternative, the one-group pretest-posttest design. And, as Stanley and Campbell noted, “the more similar the experimental and the control groups are in their recruitment, and the more this similarity is confirmed by the scores on the pretest, the more effective this control becomes.” (1963, pp. 47-48)

²Though random assignment of subjects to groups could have helped to insure group equivalence, this would be highly dependent upon the size of the sample available. Since the sample size of subjects willing to participate in the experimental group was not likely to be large (due to a highly specialized subject profile with respect to musical prerequisites and lack of existing absolute pitch ability), it was decided to recruit separate experimental and control groups from the start.
To meet these standards, subjects were recruited from the same population and every effort was made to insure that subject recruitment techniques were identical for both groups. Since the pretest could determine that the two groups were equivalent, any threats to internal validity—history, maturation, testing bias, instrument decay, statistical regression, etc.,—would affect both groups equally, and any difference between the groups would be attributable to the effects of the independent variable.

Subjects

The experimental group comprised 26 music majors from The Ohio State University School of Music, all volunteers. The total number of students enrolled in the School of Music during the training period in this study was approximately 700, over 200 of whom were graduate students. The participation of approximately 250 undergraduate upperclass students, especially juniors and seniors, was solicited through posters, classroom visitations, and personal appeals to my former students. Fourteen males and 12 females volunteered. Collectively, their mean age at the beginning of the experiment (November, 1987) was 21.9 years (median = 20, range = 19 to 33). The mean age at which they had begun formal musical training was 9.2 years (median = 9, range = 5 to 26). The mean number of years of musical training they had received was 12.7 (median = 12, range 2 to 27). The individual background data of the members of the experimental group are shown in Table 12.

³The use of volunteers was not considered a detriment to the external validity of the study because results should be generalized only to other highly motivated volunteers, typically adult musicians interested in acquiring absolute pitch.
Table 12. Background data for the individual members of the experimental group.

<table>
<thead>
<tr>
<th>Subject</th>
<th>Major</th>
<th>Rank</th>
<th>Principal</th>
<th>Age</th>
<th>Gender</th>
<th>Began</th>
<th>First Inst.</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Music Ed.</td>
<td>Junior</td>
<td>Trumpet</td>
<td>20</td>
<td>Male</td>
<td>9</td>
<td>Trumpet</td>
</tr>
<tr>
<td>B</td>
<td>Music Ed.</td>
<td>Junior</td>
<td>Voice</td>
<td>22</td>
<td>Male</td>
<td>8</td>
<td>Clarinet</td>
</tr>
<tr>
<td>C</td>
<td>Music Theory</td>
<td>Junior</td>
<td>Percussion</td>
<td>20</td>
<td>Male</td>
<td>9</td>
<td>Snare Drum</td>
</tr>
<tr>
<td>D</td>
<td>Music Ed.</td>
<td>Junior</td>
<td>Saxophone</td>
<td>20</td>
<td>Female</td>
<td>10</td>
<td>Saxophone</td>
</tr>
<tr>
<td>E</td>
<td>Music History</td>
<td>Senior</td>
<td>Bassoon</td>
<td>21</td>
<td>Female</td>
<td>6</td>
<td>Piano</td>
</tr>
<tr>
<td>F</td>
<td>Audio Eng.</td>
<td>Soph.</td>
<td>Piano</td>
<td>19</td>
<td>Male</td>
<td>14</td>
<td>Piano</td>
</tr>
<tr>
<td>G</td>
<td>Music Ed.</td>
<td>Junior</td>
<td>Voice</td>
<td>31</td>
<td>Female</td>
<td>10</td>
<td>Piano</td>
</tr>
<tr>
<td>H</td>
<td>Music Ed./Hist.</td>
<td>Senior</td>
<td>Saxophone</td>
<td>21</td>
<td>Male</td>
<td>8</td>
<td>Piano</td>
</tr>
<tr>
<td>I</td>
<td>Music Ed.</td>
<td>Junior</td>
<td>Clarinet</td>
<td>20</td>
<td>Male</td>
<td>7</td>
<td>Piano</td>
</tr>
<tr>
<td>J</td>
<td>Church Music</td>
<td>Junior</td>
<td>Voice</td>
<td>33</td>
<td>Female</td>
<td>6</td>
<td>Piano</td>
</tr>
<tr>
<td>K</td>
<td>Music Ed.</td>
<td>Senior</td>
<td>French Horn</td>
<td>21</td>
<td>Female</td>
<td>11</td>
<td>Trumpet</td>
</tr>
<tr>
<td>L</td>
<td>Music Ed./Jazz</td>
<td>Junior</td>
<td>Trumpet</td>
<td>20</td>
<td>Male</td>
<td>10</td>
<td>French Horn</td>
</tr>
<tr>
<td>M</td>
<td>Music Ed.</td>
<td>Junior</td>
<td>Voice</td>
<td>20</td>
<td>Male</td>
<td>12</td>
<td>Trumpet</td>
</tr>
<tr>
<td>N</td>
<td>Performance</td>
<td>Senior</td>
<td>Voice</td>
<td>28</td>
<td>Male</td>
<td>26</td>
<td>Voice</td>
</tr>
<tr>
<td>O</td>
<td>Music Ed.</td>
<td>Senior</td>
<td>Piano</td>
<td>31</td>
<td>Male</td>
<td>7</td>
<td>Piano</td>
</tr>
<tr>
<td>P</td>
<td>Music Ed.</td>
<td>Junior</td>
<td>French Horn</td>
<td>21</td>
<td>Female</td>
<td>9</td>
<td>Percussion</td>
</tr>
<tr>
<td>Q</td>
<td>Music Ed.</td>
<td>Senior</td>
<td>Euphonium</td>
<td>19</td>
<td>Female</td>
<td>6</td>
<td>Piano</td>
</tr>
<tr>
<td>R</td>
<td>Music Ed.</td>
<td>Soph.</td>
<td>Voice</td>
<td>21</td>
<td>Female</td>
<td>9</td>
<td>Violin</td>
</tr>
<tr>
<td>S</td>
<td>Jazz Studies</td>
<td>Senior</td>
<td>Saxophone</td>
<td>21</td>
<td>Male</td>
<td>8</td>
<td>Piano</td>
</tr>
<tr>
<td>T</td>
<td>Music Ed.</td>
<td>Junior</td>
<td>French Horn</td>
<td>20</td>
<td>Male</td>
<td>5</td>
<td>Piano</td>
</tr>
<tr>
<td>U</td>
<td>Music Ed./Th.</td>
<td>Junior</td>
<td>Piano</td>
<td>20</td>
<td>Female</td>
<td>9</td>
<td>Organ</td>
</tr>
<tr>
<td>V</td>
<td>Music Ed.</td>
<td>Junior</td>
<td>Oboe</td>
<td>20</td>
<td>Male</td>
<td>10</td>
<td>Viola</td>
</tr>
<tr>
<td>W</td>
<td>Music Ed.</td>
<td>Junior</td>
<td>Piano</td>
<td>20</td>
<td>Female</td>
<td>8</td>
<td>Perc./Piano</td>
</tr>
<tr>
<td>X</td>
<td>Music Ed.</td>
<td>Junior</td>
<td>Trumpet</td>
<td>20</td>
<td>Female</td>
<td>5</td>
<td>Piano</td>
</tr>
<tr>
<td>Y</td>
<td>Music Ed.</td>
<td>Senior</td>
<td>Percussion</td>
<td>20</td>
<td>Female</td>
<td>10</td>
<td>Voice/Perc.</td>
</tr>
<tr>
<td>Z</td>
<td>Music Ed.</td>
<td>Junior</td>
<td>Clarinet</td>
<td>20</td>
<td>Male</td>
<td>8</td>
<td>Piano</td>
</tr>
</tbody>
</table>

To be eligible to participate, subjects had to be currently enrolled music majors who had completed a 2-year aural skills sequence comprising six courses of relative pitch development. These requirements, which reduced the number of possible subjects, were implemented to eliminate the possibility of confounding the ongoing formal acquisition of relative pitch skills in their curriculum with the acquisition of absolute pitch skills. Obviously, individuals already possessing absolute pitch were not eligible for consideration as subjects. The modal responses to a questionnaire showed
that during the course of the experiment, the "average" experimental-group subject would be involved with musical activities on a daily basis, taking lessons on an instrument, performing in ensembles, and participating in both ensemble and solo performances.

A similarly formed group of 26 volunteers was recruited as a control group. Before the pretest was graded, this group contained 16 males and 10 females. Collectively, their mean age at the beginning of the experiment was 22.5 years (median = 21, range = 20 to 36). The mean age at which they had begun musical training was also 9.2 years (median = 9, range = 4 to 21). The mean number of years of their musical training was 13.3 (median = 13, range 2 to 28). Their modal responses to the questionnaire regarding their extent of musical activity were identical to those of the experimental group, and no member of either group indicated that they had any temporary or permanent hearing impairment that would impede their participation in this study. The individual background data of the members of the control group are shown in Table 13.

---

4The personal background data questionnaire for test subjects is contained in Appendix A.
Table 13. Background data for the individual members of the control group.

<table>
<thead>
<tr>
<th>Subject</th>
<th>Major</th>
<th>Rank</th>
<th>Principal</th>
<th>Age</th>
<th>Gender</th>
<th>Began</th>
<th>First Inst.</th>
</tr>
</thead>
<tbody>
<tr>
<td>AA</td>
<td>Music Ed.</td>
<td>Senior</td>
<td>Voice</td>
<td>21</td>
<td>Female</td>
<td>7</td>
<td>Piano</td>
</tr>
<tr>
<td>BB</td>
<td>Perc. Ped.</td>
<td>Grad</td>
<td>Percussion</td>
<td>27</td>
<td>Female</td>
<td>6</td>
<td>Organ</td>
</tr>
<tr>
<td>CC</td>
<td>Music Ed.</td>
<td>Junior</td>
<td>Percussion</td>
<td>20</td>
<td>Male</td>
<td>11</td>
<td>Snare Drum</td>
</tr>
<tr>
<td>DD</td>
<td>Music Ed.</td>
<td>Senior</td>
<td>Voice</td>
<td>34</td>
<td>Female</td>
<td>6</td>
<td>Piano</td>
</tr>
<tr>
<td>EE</td>
<td>Theory/Comp.</td>
<td>Senior</td>
<td>Trumpet</td>
<td>20</td>
<td>Male</td>
<td>4</td>
<td>Piano</td>
</tr>
<tr>
<td>FF</td>
<td>Audio Eng.</td>
<td>Senior</td>
<td>Trumpet</td>
<td>23</td>
<td>Male</td>
<td>9</td>
<td>Trumpet</td>
</tr>
<tr>
<td>GG</td>
<td>Music Ed.</td>
<td>Junior</td>
<td>String Bass</td>
<td>36</td>
<td>Male</td>
<td>14</td>
<td>String Bass</td>
</tr>
<tr>
<td>HH</td>
<td>Music Ed.</td>
<td>Senior</td>
<td>French Horn</td>
<td>21</td>
<td>Female</td>
<td>7</td>
<td>Piano</td>
</tr>
<tr>
<td>II</td>
<td>Performance</td>
<td>Junior</td>
<td>Violin</td>
<td>20</td>
<td>Female</td>
<td>9</td>
<td>Violin</td>
</tr>
<tr>
<td>KK</td>
<td>Audio Eng.</td>
<td>Senior</td>
<td>Voice</td>
<td>21</td>
<td>Male</td>
<td>5</td>
<td>Voice</td>
</tr>
<tr>
<td>LL</td>
<td>Music Ed.</td>
<td>Senior</td>
<td>Voice</td>
<td>22</td>
<td>Female</td>
<td>6</td>
<td>Piano</td>
</tr>
<tr>
<td>MM</td>
<td>Music Ed.</td>
<td>Senior</td>
<td>Saxophone</td>
<td>21</td>
<td>Male</td>
<td>10</td>
<td>Saxophone</td>
</tr>
<tr>
<td>NN</td>
<td>Music Ed.</td>
<td>Junior</td>
<td>Trombone</td>
<td>20</td>
<td>Male</td>
<td>11</td>
<td>Piano</td>
</tr>
<tr>
<td>OO</td>
<td>Music Ed./Perf.</td>
<td>Junior</td>
<td>Piano/Viola</td>
<td>20</td>
<td>Female</td>
<td>7</td>
<td>Piano</td>
</tr>
<tr>
<td>PP</td>
<td>Music Ed.</td>
<td>Junior</td>
<td>Euphonium</td>
<td>21</td>
<td>Male</td>
<td>12</td>
<td>Trumpet</td>
</tr>
<tr>
<td>QQ</td>
<td>Audio Eng.</td>
<td>Junior</td>
<td>Voice</td>
<td>24</td>
<td>Male</td>
<td>10</td>
<td>Guitar</td>
</tr>
<tr>
<td>RR</td>
<td>Ind. Stud. Mus.</td>
<td>Junior</td>
<td>Piano</td>
<td>21</td>
<td>Female</td>
<td>5</td>
<td>Piano</td>
</tr>
<tr>
<td>SS</td>
<td>Music Ed.</td>
<td>Junior</td>
<td>Violin</td>
<td>20</td>
<td>Male</td>
<td>7</td>
<td>Piano</td>
</tr>
<tr>
<td>TT</td>
<td>Performance</td>
<td>Grad</td>
<td>Saxophone</td>
<td>22</td>
<td>Male</td>
<td>10</td>
<td>Saxophone</td>
</tr>
<tr>
<td>UU</td>
<td>Music Ed.</td>
<td>Junior</td>
<td>Voice</td>
<td>20</td>
<td>Female</td>
<td>8</td>
<td>Piano</td>
</tr>
<tr>
<td>VV</td>
<td>Music Ed.</td>
<td>Junior</td>
<td>Euphonium</td>
<td>22</td>
<td>Female</td>
<td>8</td>
<td>Guitar</td>
</tr>
<tr>
<td>WW</td>
<td>Music Ed.</td>
<td>Junior</td>
<td>French Horn</td>
<td>20</td>
<td>Male</td>
<td>10</td>
<td>Trumpet</td>
</tr>
<tr>
<td>XX</td>
<td>Ind. Stud. Mus.</td>
<td>Senior</td>
<td>Piano</td>
<td>23</td>
<td>Male</td>
<td>21</td>
<td>Piano</td>
</tr>
<tr>
<td>YY</td>
<td>Jazz Studies</td>
<td>Senior</td>
<td>String Bass</td>
<td>22</td>
<td>Male</td>
<td>13</td>
<td>String Bass</td>
</tr>
<tr>
<td>ZZ</td>
<td>Music Ed.</td>
<td>Senior</td>
<td>Trumpet</td>
<td>23</td>
<td>Male</td>
<td>11</td>
<td>Trumpet</td>
</tr>
</tbody>
</table>

Only members of the experimental group received absolute pitch training. Two of the members of the control group, subjects OO and TT, indicated on their questionnaires that they believed they already possessed absolute pitch, but still wished to participate in the experiment. Remembering that the first conclusion made in Lundin and Allen’s (1962) training experiment in absolute pitch was that the 2 subjects who claimed to possess absolute pitch were mistaken,\(^5\) judgment regarding the appropriateness of these subjects’ participation was withheld pending their

\(^5\)Lundin and Allen’s subjects scored 54% and 42% correct responses to a 24-tone pretest of the chromatic pitches ranging from C\(_4\) to B\(_5\).
taking of the pretest. If they did in fact possess absolute pitch, it was felt that they would not be representative of the population recruited for participation in the experiment. Analysis of their pretest scores indicated levels of ability far superior to those of the other subjects, and their data, shown in Appendices B and C, were not included with those of the control group. Subjects OO and TT were still paid for their participation, and took the posttest as well. This was not deemed detrimental to the study—the other subjects were unaware of their disqualification—and, in fact, measurement of two absolute pitch possessors provided an opportunity to gauge the reliability and the validity of the test instrument.

The mean age at the beginning of the experiment of the new 15-male, 9-female control group was 22.7 years (median = 21, range = 20 to 36). The mean age at which they had begun musical training was 9.4 years (median = 9.5, range = 4 to 21). The mean number of years of their musical training was 13.3 (median = 13, range 2 to 28). Their modal responses to the questionnaire regarding their extent of musical activity were identical to those of the experimental group.

Description of the Pretest and Posttest

A test instrument was needed to measure the absolute pitch-naming skills of the experimental and control groups prior to training, because these two groups were not randomly selected. To measure passive absolute pitch, the test would have to involve a categorization (identification) task, whereby the subject would indicate the pitch class of a musical tone without the aid of an external reference tone. This working definition of "passive" absolute

---

6Subjects OO and TT scored 85% and 100% correct, respectively, on the pretest.
pitch involves only one aspect of the broader, widely held definition of absolute pitch as "the ability to identify the frequency or musical name of a specific tone, or, conversely, the ability to produce some designated frequency, frequency level, or pitch without comparing the tone to any objective reference tone (i.e., without using relative pitch)." (Ward & Burns, 1982, p. 431)

No standardized test of absolute pitch ability was found in the course of reviewing the literature on absolute pitch, not even in that segment of the literature concerning training. In fact, the only absolute pitch test that had a history of repeated usage was that developed by Dr. W. Dixon Ward, an otolaryngologist at the University of Minnesota. Usage of his "Test 1" can be traced back over 35 years, and descriptions of it are contained in Ward (1953), Ward (1963b), Waters (1980), and Ward & Burns (1982). Ward & Burns wrote, "we routinely use this test to screen for possessors of AP [absolute pitch] in groups of listeners." (p. 440) Ward's test comprises 100 items arranged in five blocks of 20, and each item is a 2-second sine tone followed by 20 seconds of silence. However, Ward's test uses only 10 of the 12 possible chromata, none of which are presented in an octave other than the original. It allows identical tones to appear in immediate succession, and fails to relate the manner of subject response to any type of familiar musical behavior. Furthermore, various accounts state that at times the tones were presented to the left ear only, and sometimes in an anechoic chamber. While the test developed for this study was modeled after the best work available (Ward's), certain features were introduced to extend its usefulness and musicality.

The test developed and utilized in this experiment contains 120 items, 10 from each of the 12 pitch classes. The range extends from C2 to B5 (the two
octaves above and below “middle C”); this is the range required for practice in Burge’s Color Hearing Technique Exercises. Items are presented in five blocks of 24 tones each. Each tone in the sequence varies randomly within a small range of loudness, timbre and duration. Each test item consists of a 2-second piano tone, followed by 3 seconds of silence to allow subjects to indicate their answers in writing. This interstimulus time interval was deemed to be just sufficient for subjects to record their answers, if they recognized the pitches immediately. The answer sheets were collected as soon as the last item was presented so there was no time for the subjects to check over their responses. No physical attempt was made to erase the subjects’ short term pitch memories. Subjects were instructed not to attempt to utilize relative pitch during testing, and were not allowed to hum, sing, whistle or talk once the test had begun. No correction was provided during testing. Item numbers are given only at the beginning of each new block, as opposed to between each item, and each block is separated by a pause of 20 seconds.

Tones were presented in a quasi-random order, with the only restriction being that each pitch class appeared twice in each block, but not consecutively, even at the boundaries of adjacent blocks.\(^7\) The octave of each

\(^7\)Rather than trust computer randomization schemes, a collection of twenty-four playing cards was assembled for the randomization of the pitch classes. Two of every face were included, excluding the King. The pitch class labels assigned to equate card face to pitch class number were as follows; Queen = 0, Ace = 1, 2 = 2, 3 = 3, 4 = 4, 5 = 5, 6 = 6, 7 = 7, 8 = 8, 9 = 9, 10 = 10, Jack = 11. Cards were randomly drawn without replacement to determine the pitch class order of the test. No significance was attached to the playing cards’ suits. When two identical pitch classes were successively drawn, the draw was deemed invalid, and the card was reshuffled back into the remainder of the deck.
tone was quasi-randomly selected as well, as each block contains six tones 
from each of the four possible octaves.8

Listeners were not given information regarding the makeup of the test 
items other than the total number of test items (the block length and number 
of blocks) and the response time duration. The same test was given for both 
the pretest and posttest measures, but subjects were not informed of this. 
They had no knowledge of their test performance, and even if they had it is 
doubtful that any of them could have remembered any of the test items after 
34 weeks.

The test stimuli consisted of a digital recording of a 9-foot Steinway & 
Sons concert grand piano.9 The instrument was played by a professional 
pianist, Dr. Thomas Wells of The Ohio State University School of Music 
faculty. The stimuli were performed with the aid of a “click-track” to help the 
pianist monitor the temporal accuracy of his performance.10 Prior to the 
recording, the piano received a professional 12-to-the-octave equal-tempered

---

8For the randomization of octaves a collection of twenty-four playing cards 
was assembled, equally representing six of each suit. Assignments were made 
as follows; Spades = 2nd octave, Hearts = 3rd octave, Clubs = 4th octave, 
Diamonds = 5th octave. Cards were randomly drawn without replacement. 
No significance was attached to face value and no restrictions regarding 
succession were invoked. This procedure produced five blocks of twenty-four 
pitches. These blocks were quasi-randomly ordered; the first pitch class of a 
block could not be the same as the last pitch of the previous block. 

9It has been suggested that possession of “absolute piano” may not quite be the 
same as possession of “absolute pitch” (Ward, 1963a, p. 17, and Burns & Ward, 
1982, p. 437). For the sake of ecological validity, and since the majority of 
literature on absolute pitch is based on piano-tune studies, real musical 
stimuli—as opposed to sinusoidal stimuli—were used. Indeed, several classic 
studies have found the timbre of the piano the easiest of all timbres for 
absolute pitch possessors to identify, regardless of ability, principal 
instrument, or background (von Kries, 1892; Abraham, 1901; Baird, 1917; 
Mull, 1925; Petran, 1932).

10A copy of the score may be found in Appendix A.
tuning by Mr. Paul Schopis, the piano technician for the School of Music. The instrument was in-tune with itself, as well as to the A4 = 440 Hz tuning standard.

A stereo recording was made on the concert stage of Weigel Hall, a 720-seat concert hall of The Ohio State University. Mr. Richard Chitty, Director of the O.S.U. Audio Recording program, personally performed all recording. Equipment utilized in the recording process consisted of a coincident pair of Crown PCC-160 phase-coherent cardioid recording microphones and a Sony PCM-601ESD digital audio processor interfaced with a Sony SL-HF450 Super Beta hi-fi stereo video cassette recorder. The sampling frequency used was 44.1 kHz, and the 16-bit linear quantizing mode of this system produced two channel output with a frequency response from 5 Hz-20 kHz ± 0.5 dB. The harmonic distortion was less than 0.005%, the dynamic range was greater than 90 dB, the channel separation was greater than 80 dB, and the wow and flutter were below measurable limits. While in digital format, data were edited and stored on TDK HD Pro Super Avilyn high definition professional master quality video cassette tape.

Listeners were asked to indicate on a numbered answer sheet provided the pitch class name of each tone presented; they were not asked to identify the proper octave, but were told to answer every item. Listeners were instructed not to whistle or hum, and not to converse among themselves once the test had begun. They were also instructed not to attempt to use their relative pitch ability (i.e. to relate any tone to any previous tone). A cassette audio tape of the test directions was played for them and they were instructed
to read along with it. At no time were they given feedback about the correctness of their responses.

Testing took place in the Sound Synthesis Studios of The Ohio State University. This facility has excellent acoustic qualities, and amply permits group testing. The walls have special sound-deadening panels on them. During testing, the room was additionally protected from outside noises by means of two independent sets of closed doors. Listeners were presented the test stimuli in a open sound field and no extraneous sounds were audible. During playback of the test, the signal was directed through a Quantum mixing console to Phase Linear Model 200 Series Two Audio Standard power amplifiers driving JBL Model 4331A studio reference monitors.

Training

The optimal treatment period available for subject training was three consecutive academic quarters, the 1987-88 academic year. A longer training schedule was not feasible because summer quarter undergraduate music major enrollment is prohibitively low. Even a three-quarter training period was feared to be, in terms of subject mortality, quite long. Incentives were given to encourage experimental group subjects not to drop out of the study.

The fact that all experimental group subjects were volunteers, and were therefore obviously interested in trying to acquire absolute pitch skills, was of great help. Furthermore, only individuals who agreed from the outset to commit themselves to a 25 minutes-a-day training regimen for the entire training period were accepted as participants. Prospective subjects who

---

11A copy of the test directions is contained in Appendix A.
12A copy of the subject consent form required by The Ohio State University for experimentation involving human subjects is contained in Appendix A.
considered themselves less than fully committed to completing the project were counseled out of the experiment. Regardless, the expected level of participation represented a significant time commitment for an uncompensated individual to make. Implementation of a reward system for participation which encouraged progress provided additional incentive for subjects to continue in the experiment for the entire three quarters.

All subjects, both in the experimental and control groups, were paid. Compensation in the amount of five dollars, paid at the posttest, was made to all subjects for merely taking both the pretest and the posttest. Members of the experimental group received the use of the training materials at no cost,\textsuperscript{13} and those who trained for the entire year were entitled to keep their handbooks. Additionally, subjects knew that those members of the experimental group who reached a predetermined training goal, completion of Exercise 15, would earn equal shares of a $1,250 sum, payable in certified check immediately after the experiment's conclusion. These incentives proved sufficient to attract subjects at the outset, while not providing a situation where experimental group subjects were merely paid to participate. The amount of compensation was clearly insufficient for those whose sole motivation was monetary.

The 1987-88 academic year began on 23 September, 1987. Recruitment of subjects began immediately, and lasted for 3 weeks. Subjects were

\textsuperscript{13}These rather substantial expenses were borne by the experimenter. This was done to avoid the appearance of any impropriety. If the effect of the training was found to be negligible, the experimenter could freely report so without any duty to the author of the training technique, or to his publisher. Conversely, if the training techniques were found to produce a substantial effect, the experimenter could not be accused of receiving favors from the author.
encouraged to enter the study with a partner, because training was to take place in pairs. The fact that most of the subjects provided their own partners (who in most cases were already friends) was considered beneficial to their odds of staying in the experiment. Subjects who lacked a partner at the outset were provided an opportunity to obtain one from a list of people who expressed an interest in participating, but who also lacked a partner. Partners were required to play the same instrument, or at least be able be able to train on a like instrument.\textsuperscript{14}

On 14 October, 1987 the pretest was administered to both the experimental and control groups, and training began. The first event in the training involved the playing of the \textit{Perfect Pitch Seminar Tapes} (\textit{Workshop} and \textit{Master Class} tapes) for the experimental group. These tapes introduce and elaborate on the material contained in the handbook. Next, the handbooks were distributed, and subjects were instructed to begin training at their earliest possible convenience. Later, the \textit{Color Hearing “Help Me” Tapes} and the \textit{Perfect Pitch Solo Tape—Private Lesson Tape} were played for the experimental group, again in strict accordance to their directions. These tapes do not contain “drills,” but serve as a supplement to the handbook. All tapes were available to the subjects for their individual use throughout the experiment.

The choice of instrument on which to train was left to the subjects. They were directed to consult their handbooks in making their decisions. Inspection of the detailed practice logs revealed that all eventually chose the

\textsuperscript{14}For example, a pianist and a vocalist were permitted to be partners because both could train using the piano.
piano,\textsuperscript{15} in accordance with the handbook, which states "for convenience in drilling, the piano is probably the best instrument to use, and therefore all exercises are geared to the keyboard." (Burge, 1973, p. 40)

It is important to note that the experimenter never presented himself as a "teacher" of the training technique. In fact, the objective nature of the experimenter's role was stressed to the subjects. In an effort to minimize experimenter bias, the experimenter has never practiced the training techniques either prior to the experiment, or during its course. No attempt was made to hold training "classes," and every effort was made to provide a training situation most like the one normally encountered by one who would train using this technique outside of this experiment. Subjects agreed to follow the directions of the method itself, and any subjects posing questions were referred to either the manual or the tapes.

\textsuperscript{15}An example of one of the weekly practice logs is contained in Appendix A. These logs provided the author with detailed information regarding the date and location of each practice session, its exact starting and ending times, the number of the exercise practiced, the practice instrument used, and whether the session was a team or solo effort.
CHAPTER V

RESULTS AND DISCUSSION

Interpretation of Test Results

The test instrumentation for this experiment conformed to the type of absolute pitch test most commonly used—the categorization of a series of piano tones for the subject to identify by pitch name.\(^1\) Likewise, pretest and posttest data for each subject were scored according to the three most commonly used measures: the total number correct, the mean absolute error in semitones, and the amount of information transmitted in terms of information theory.\(^2\)

Of the three measures, the total number (or proportion) correct, though the most easily understood and easiest to determine, is nonetheless the least elegant. While directly measuring accuracy level, it does not gauge variability of responses. Consider, for example, the following set of data derived from two hypothetical subjects. The first subject answers in a completely random fashion, obtaining a chance number of correct responses and an equal distribution of all possible errors. The second subject makes no correct

---

\(^1\)While this categorization procedure might theoretically discriminate against subjects with exceptional pitch memory but who are inexperienced with pitch names, no problem was encountered because the subjects were all highly trained musicians.

\(^2\)Ward 1963a, p. 18.
responses, but always errs by a semitone in the same direction. According to the total number of correct responses, the first subject would fare better than the second; however, the lower variance of the latter subject's responses clearly indicates some level of skill that this measure overlooks. A subject similar to the hypothetical second one mentioned here is cited by Weinert (1929), and has been observed directly by this writer. Measurements of this nature would also discriminate against possessors of absolute pitch who have reported judgments that have shifted a semitone (or more) higher with age (Triepel, 1934; Vernon, 1977; Gabrielsson, 1981).

The second measure of absolute pitch ability used in this study was the calculation of the mean absolute error in semitones. The term "absolute" is used in this context in its mathematical sense, implying absolute value (where only magnitude is important, as both positive and negative directions are made positive). If errors were not measured in this way, the mean error of a hypothetical subject whose sum of errors in the negative direction equalled the sum of errors in the positive direction would be zero, even though one could have, in fact, made no correct responses. Similarly, consider the comparison of two more hypothetical subjects, in this case with identical mean absolute error scores. The first might make all errors of a small magnitude, while correctly identifying none. The second subject might

---

3 A former student of mine had, at a very early age, learned the pitch names of his reliable, yet inaccurate, absolute pitch ability according to his performance instrument—unfortunately in this case—a B♭ trumpet. Also see footnote 6, infra, p. 154.

4 Changes in the elasticity of the basilar membrane with age, which would cause a shift in the point of excitation for any given pitch toward the oval window and away from the helicotrema, have been suggested as an explanation for this phenomenon (Ward & Burns, 1982; Campbell & Greated, 1987).
make several correct judgments, yet be prone to errors of a large magnitude. The observation that these two subjects are "equal" seems false.

The final, and most appropriate, measure for the analysis of the absolute pitch data is the calculation of the amount of information transmitted in terms of information theory. Procedures defining this technique were described in the early 1950s by Garner & Hake (1951), and were first applied to absolute pitch by Pollack (1952). An excellent summary of the technique, and its appropriateness, was provided by Ward (1963a):

The "information" in a signal is given by the logarithm, to the base 2, of the number of alternatives from which it might have been selected. In essence, this is the minimum number of times that one would have to divide the stimulus population into two groups in order to be sure of finally isolating the desired one. Thus, if there are eight possible stimuli, one must first divide them into two groups of four, each of these into groups of two, and each of these into groups of one. Since there were three steps, there is said to be three "bits" of information in the signal.

If the listener always identifies the signal correctly, then the information transmitted is the same as that of the signal. The more errors he makes, however, the lower will be the amount of information transmitted. Consistency is rewarded as much as accuracy in this scheme. . . . This procedure is ideal for categorization studies of AP [absolute pitch]. One defines the possible stimuli to the listener, presents each an equal number of times while the listener attempts identification, and tabulates the responses. From this, the information transmitted, i.e., the information in the response, \( I_r \), can be calculated. (pp. 36-37).

Extending Ward's analogy, the amount of information contained in a correct response among a choice of two possible stimuli would yield 1 bit of information. A correct response among a choice of four possible stimuli would yield 2 bits of information. As noted earlier, during testing, subjects were asked to respond only with a pitch class label; no octave identification
was required. There being 12 discrete response categories (chromata), a correct response to any test item would transmit \( \log_2 (12) = 3.585 \) bits of information.

A question arises when considering the computation of the amount of information contained in incorrect responses. Carroll (1975) devised a procedure allowing specifically for the calculation of information according to chroma, separate from octave information, and indexed “the amount of information transmitted by each response relative to its deviation from the correct stimulus category.” (p. 24) According to Carroll:

A response with an error of one semitone in either direction has the effect of carrying the amount of information transmitted if the 12 semitones of an octave were divided into 4 sets of 3 semitones; that is, the response “correctly” assigns a stimulus to one of these 4 sets. Similarly, a response with an error of 2 semits from the stimulus assigns the stimulus to one of 2.4 sets of 5 semitones each. Generalizing for any given amount of response error (e being the absolute magnitude of the error in semits [semitones]), we find the amount of chroma information (\( I_C \)) as

\[
I_C = \log_2 \left[ \frac{12}{(2e + 1)} \right].
\]

(p. 25)

Carroll’s information theory formulae were used because the most commonly cited example of “success” in attaining absolute pitch skills as an adult, though far from being scientific, is that of Brady (1970), and the only independent measurement and confirmation of this was conducted by Carroll (1975). This decision also allowed for comparisons to be made between the

---

5This decision to measure chroma identification, exclusive of octave identification, is consistent with the training method under consideration, and additionally eliminated the problematic question of how to account for octave errors.
results of this study, and those of Carroll's (which included 4 native possessors of absolute pitch besides Brady).

Discussion of Results Obtained from the Complete Experimental and Control Groups

The 120-item pretests and posttests were scored according to the mean number of correct responses, the mean absolute error in semitones, and the amount of information transmitted in bits. No evidence of constant errors, which might indicate either the attempt to use a highly accurate sense of relative pitch, or perhaps a "mis-tuned" subject, was found.\(^6\)

The data obtained from the measurement of the mean number of correct responses for the experimental and control groups are contained in Table 14.

**Table 14.** Mean number of correct responses (120 possible), standard deviation, and number of observations on the pretest and the posttest for the experimental and control groups.

<table>
<thead>
<tr>
<th>Experimenta Group</th>
<th>Control Group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pretest</td>
</tr>
<tr>
<td>Mean</td>
<td>16.538</td>
</tr>
<tr>
<td>SD</td>
<td>12.452</td>
</tr>
<tr>
<td>N</td>
<td>26</td>
</tr>
</tbody>
</table>

To help convey the relationship between these two sets of data, the group means contained in Table 14 are shown in Figure 6.

\(^6\)Ward 1963a, p. 18.
Figure 6. The effect of training on the mean number of correct responses for the experimental and control groups.

A parametric test of significance was used to determine whether the obtained differences between the two groups exceeds the level of fluctuation to be expected in cases of no true difference for samples of that size. Student’s $t$ tests for independent samples were used to determine whether the differences among group means were statistically significant at the .05 probability level. The results indicate that before training the pretest means of the two groups were not significantly different ($t = 1.317$, $df = 48$, $p = .194$). Following training, the posttest means of the two groups were significantly different ($t = 2.642$, $df = 48$, $p = .011$).
The data obtained from the measurement of the mean absolute error for the experimental and control groups are contained in Table 15.

**Table 15.** Mean absolute error (semitones), standard deviation, and number of observations on the pretest and the posttest for the experimental and control groups.

<table>
<thead>
<tr>
<th></th>
<th>Experimental Group</th>
<th>Control Group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pretest</td>
<td>Posttest</td>
</tr>
<tr>
<td>Mean</td>
<td>2.730</td>
<td>2.365</td>
</tr>
<tr>
<td>SD</td>
<td>0.565</td>
<td>0.812</td>
</tr>
<tr>
<td>N</td>
<td>26</td>
<td>26</td>
</tr>
</tbody>
</table>

To help convey the relationship between the two groups, the group means contained in Table 15 are shown in Figure 7.
Figure 7. The effect of training on the mean absolute error for the experimental and control groups.

Student’s t tests for independent samples were used to determine whether the differences among group means were statistically significant at the .05 probability level. The results indicate that before training the pretest means of the two groups were not significantly different ($t = -1.100, df = 48, p = .277$). Following training, the posttest means of the two groups were significantly different ($t = -3.118, df = 48, p = .003$).

The data obtained from the measurement of the amount of information transmitted for the experimental and control groups are contained in Table 16.
Table 16. Mean amount of information transmitted (bits), standard deviation, and number of observations on the pretest and the posttest for the experimental and control groups.

<table>
<thead>
<tr>
<th></th>
<th>Experimental Group</th>
<th>Control Group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pretest</td>
<td>Posttest</td>
</tr>
<tr>
<td>Mean</td>
<td>1.243</td>
<td>1.503</td>
</tr>
<tr>
<td>SD</td>
<td>0.380</td>
<td>0.601</td>
</tr>
<tr>
<td>N</td>
<td>26</td>
<td>26</td>
</tr>
</tbody>
</table>

It is interesting to note that while the control group subjects had a higher number of correct responses on the posttest (Table 14), the size of their mean absolute error actually increased slightly (Table 15). While these figures may seem to be inconsistent with each other, they are both, if fact, true, and the mean amount of information transmitted by the control group reconciles these discrepancies (Table 16). Their level of ability, while remaining very stable, did fall slightly.

To help convey the relationship between the two groups, the group means contained in Table 16 are shown in Figure 8.
Figure 8. The effect of training on the mean amount of information transmitted for the experimental and control groups.

Student's t tests for independent samples were used to determine whether the differences among group means were statistically significant at the .05 probability level. The results indicate that before training the pretest means of the two groups were not significantly different ($t = 1.156$, $df = 48$, $p = .254$). Following training, the posttest means of the two groups were significantly different ($t = 2.950$, $df = 48$, $p = .005$).

It is appropriate at this point in the discussion of the results to consider a way in which the data were not, and should not, be analyzed. In order to determine the effectiveness of the training, one might consider directly comparing the pretest and posttest scores within groups, and comparing the
amount of improvement of one group relative to that of the other. This
method is unsuitable, because the real issue is whether, following treatment,
the experimental group is better than the control group. Moreover, the direct
calculation of difference scores, or gain scores (the posttest score minus the
pretest score), for the subjects introduces the inherent problem that all
students do not have the same amount of potential gain available to them.
Gay (1981) effectively summarizes this problem with this simple example:

On a 100-item test, who is better, a student who goes from a pretest
score of 80 to a posttest score of 99 (a gain of 19), or a student who goes
from a pretest score of 20 to a posttest score of 70 (a gain of 50)? (p. 229)

Thus, the only appropriate comparison is of the posttest scores; pretest
comparisons exist merely to establish the similarity between the experimental
and control groups because they were not the products of random selection.

To determine whether the posttest superiority of the experimental
group was the result of increased accuracy to one or several chromata, all
pretest and posttest responses were separated according to pitch class and then
analyzed. Group means were tabulated and plotted as a function of pitch
class, and the results are shown in Figures 9 and 10.
Figure 9. Mean number of correct responses, mean absolute error, and mean amount of information transmitted as a function of pitch class for the subjects in the experimental group.
Figure 10. Mean number of correct responses, mean absolute error, and mean amount of information transmitted as a function of pitch class for the subjects in the control group.
Figure 9 clearly indicates that the improved posttest mean scores of the experimental group were not the result of improvement on just a few chromata. While the accuracy of the experimental group was the greatest for the pitch classes F#, A, and C following training, all except B♭ show improvement from the pretest.

Figure 10 indicates that the subjects in the control group were very consistent in their responses from pretest to posttest. Though separated by 34 weeks, the patterns of their responses were virtually identical. The plot of the mean number of correct responses as a function of pitch class revealed the tendency of these subjects to name members of the C diatonic collection correctly more often than the nondiatonic ones, a phenomenon also evidenced by the pretest plot of the experimental group. But, this phenomenon could also indicate only that the diatonic responses were offered more often, perhaps the result of "lazy" guessing. If this were the case, mistakes would be of random size; indeed, this inference was confirmed in the plot of the mean absolute error by pitch class. The calculation of the mean amount of information transmitted by pitch class, the lowest graph, reconciles the upper two graphs for both groups.\(^7\)

All pretest and posttest responses were also separated according to the size of their deviation from the correct response in semitones and then analyzed. Group means were tabulated and plotted as a function of error interval and absolute error interval; the results are shown in Figures 11 and 13 for the experimental group, and Figures in 12 and 14 for the control group.

\(^7\)Individual graphs of the number of correct responses, the mean absolute error, and the amount of information transmitted as a function of pitch class for all subjects are shown in Appendix B.
Figure 11. Mean number of responses as a function of error interval for the subjects in the experimental group.

Figure 12. Mean number of responses as a function of error interval for the subjects in the control group.
Figure 13. Mean number of responses as a function of absolute error interval for the subjects in the experimental group.

Figure 14. Mean number of responses as a function of absolute error interval for the subjects in the control group.
Figures 11 and 13 indicate that the increased accuracy of the subjects in the experimental group was the result of their ability to make more correct judgments and to reduce the size of the error in their incorrect judgments. Of course, these two go hand in hand because the sum of all responses on any test must equal 120, but the increase in correct judgments appears to have been accompanied by a reduction of errors 4 and 5 semitones above and below the target (major third, and perfect fourth). Interestingly, the number of tritone errors remained constant.8

Figures 12 and 14 indicate the distribution of errors committed by the control group on the posttest was not only similar to that of the pretest, but still almost random.9

The posttest scores of the experimental group rose more than those of the control group, and so did their standard deviations (see Tables 14, 15, and 16). Clearly, during the course of the experiment the composition of experimental group became more heterogeneous, while that of the control group became more homogeneous. Figures 15, 16, 17, 18, 19, and 20 indicate the dispersion of the pretest and posttest scores for the individual subjects in the experimental group and the control group.

---

8Individual graphs of the number of responses as a function of error interval, and the number of responses as a function of absolute error interval for all subjects are shown in Appendix C.
9A random distribution of the 120 responses as a function of error interval would produce a histogram of 10 responses in every category. A random distribution of the 120 responses as a function of absolute error interval would produce a histogram of 10 responses in the 0 semitones and 6 semitone error categories, and 20 responses in categories 1-5 because correct responses and tritone errors have no meaning above or below the target.
Figure 15. Individual scores of the number of correct responses for the subjects in the experimental group.

Figure 16. Individual scores of the number of correct responses for the subjects in the control group.
Figure 17. Individual scores of the mean absolute error for the subjects in the experimental group.

Figure 18. Individual scores of the mean absolute error for the subjects in the control group.
Figure 19. Individual scores of the amount of information transmitted for the subjects in the experimental group.

Figure 20. Individual scores of the amount of information transmitted for the subjects in the control group.
Divisions of the Experimental Group

The members of the experimental group trained a combined total of 52,387 minutes, distributed over a total of 2,274 training sessions. The mean number of training sessions completed per subject was 87.5 (median = 70, range = 2 to 230), and the mean number of minutes of training per subject was 2,015 (median = 1,594.5, range = 40 to 6,600). The relatively large ranges exhibited here indicate that the amount of training received per subject was not uniform. Though beyond the control of the experimenter, this lack of uniformity was confirmed by inspection of individual subjects' training logs. The relationship of the means relative to the medians indicates that both distributions are positively skewed ($Sk = 0.724$ and $1.092$, respectively), which implies that while some subjects practiced quite extensively, many did not.

Thus, the experimental group experienced some degree of mortality; that is, cases where the effect of the treatment appears diluted due to some form of subject attrition. Certainly, some level of mortality could be expected for a study of this length. As it turned out, the loss was not attributable to a lack of validity of the treatment per se, but rather to differential attendance at training sessions. On this matter, Campbell and Stanley (1963) have written that:

Mortality, lost cases, and cases on which only partial data are available, are troublesome to handle, and are commonly swept under the rug. Typically, experiments on teaching methods are spread out over days, weeks, or months. If the pretests and posttests are given in the classrooms from which experimental group and control group are drawn, and if the experimental condition requires attendance at certain

---

10 Individual training log summaries are contained in Appendix D, and the comments of the individual experimental group subjects, many of which concern mortality, are contained in Appendix E.
sessions, while the control condition does not, then the differential attendance on the three occasions (pretest, treatment, posttest) produces “mortality” which can introduce subtle sample biases. If, of those initially designated as experimental group participants, one eliminates those who fail to show up for experimental sessions, then one selectively shrinks the experimental group in a way not comparably done in the control group, biasing the experimental group in the direction of the conscientious and healthy. The preferred mode of treatment, while not usually employed, would seem to be to use all of the selected experimental group and control students who completed both pretest and posttest, including those in the experimental group who failed to get the X [treatment]. This procedure obviously attenuates the apparent effect of the X, but it avoids the sampling bias. (p. 15-16)

This was precisely the way in which the data were treated in this study. In an attempt to represent the effect of the training more accurately, the experimental group data were divided into groups according to the highest difficulty level that they practiced in training and the total amount of training they received. Division was not based in any way on subjects’ performance on either the pretest or the posttest. Three mutually exclusive groups were formed as follows: Group A comprised those subjects who attained the predetermined training criterion (advancement to Lesson 15); Group B comprised those subjects who devoted a considerable amount of time to their training, yet failed to reach the criterion; and Group C comprised those subjects who failed to devote a substantial amount of time to their training. A rank-ordering of the total durations of subjects’ practice times revealed a large gap centered on 1,200 minutes of practice time, which provided the boundary for dividing Groups B and C. The following discussion will treat each group individually, in order of increasing levels of participation.
Discussion of Results Obtained from Group C

Group C was composed of 9 subjects: Subjects B, E, F, G, I, K, L, V and Z. None trained more than 1,200 minutes, in fact, not one trained even 1,000 minutes. They trained a combined total of 4,503 minutes, distributed over a total of 236 training sessions. The mean number of training sessions completed per subject was 26.2 (median = 34, range = 2 to 43), and the mean number of minutes of training per subject was 500.3 (median = 483, range = 40 to 948). While these nine subjects accounted for 34.6% of the experimental group population, their efforts amounted to only 10.4% of all practice sessions and 8.6% of all practice time. They were, in effect, “drop-outs”; each had discontinued training even before the halfway point of the experiment. The mean highest difficulty level attained by the subjects in Group C was Lesson 4. The data obtained from the measurement of the mean number of correct responses, the mean absolute error, and the amount of information transmitted for Group C and the control group are contained in Table 17.
Table 17. Mean number of correct responses (120 possible), mean absolute error (semitones), mean amount of information transmitted (bits), their standard deviations, and number of observations on the pretest and the posttest for Group C and the control group.

<table>
<thead>
<tr>
<th></th>
<th>Group C</th>
<th>Control Group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pretest</td>
<td>Posttest</td>
</tr>
<tr>
<td>Number Correct</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>14.778</td>
<td>16.111</td>
</tr>
<tr>
<td>SD</td>
<td>10.220</td>
<td>8.937</td>
</tr>
<tr>
<td>N</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>Mean Absolute Error</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>2.827</td>
<td>2.646</td>
</tr>
<tr>
<td>SD</td>
<td>0.505</td>
<td>0.515</td>
</tr>
<tr>
<td>N</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>Amount of Information Transmitted</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>1.175</td>
<td>1.275</td>
</tr>
<tr>
<td>SD</td>
<td>0.337</td>
<td>0.317</td>
</tr>
<tr>
<td>N</td>
<td>9</td>
<td>9</td>
</tr>
</tbody>
</table>

To help convey the relationship between the two groups, the group means contained in Table 17 are shown in Figure 21.
Figure 21. Mean number correct, mean absolute error, and mean amount of information transmitted for subjects in Group C and the control group.
Student's $t$ tests for independent samples were used to determine whether the differences among posttest group means were statistically significant at the .05 probability level. While Group C's posttest means indicate some improvement occurred as a result of training, these changes were not significant for any measure—the total number correct ($t = 1.344, df = 31, p = .189$), the mean absolute error ($t = -1.930, df = 31, p = .063$), or the mean amount of information transmitted ($t = 1.752, df = 31, p = .090$). Group C's failure to experience significant improvement over the control group is hardly surprising, inasmuch as membership in Group C was singularly distinguished on the basis of a limited amount of training.

The pretest and posttest responses of the nine subjects comprising Group C were separated according to pitch class and then analyzed, as had been done previously with the full experimental and control groups. Group means were tabulated and plotted as a function of pitch class, and the results are shown in Figure 22.
Figure 22. Mean number of correct responses, mean absolute error, and mean amount of information transmitted as a function of pitch class for the subjects in Group C.
Figure 22 indicates that the group responses on the posttest were quite similar to those on the pretest, except for the slight improvement on the chromata F, F#, and G, accompanied by substantial improvement on the chroma D. This is quite unexplainable since practice on this pitch class was to have been undertaken in Lesson 11, and no Group C subject advanced past Lesson 5.

All pretest and posttest responses for the subjects in Group C were separated according to the size of their deviation from the correct response in semitones and then analyzed. Group means were tabulated and plotted as a function of error interval and absolute error interval, and the results are shown in Figures 23 and 24.
Figure 23. Mean number of responses as a function of error interval for the subjects in Group C.

Figure 24. Mean number of responses as a function of absolute error interval for the subjects in Group C.
The frequency distributions in Figures 23 and 24 indicate that, from the pretest to the posttest, the changes in the mean number of responses as a function of error interval and absolute error interval for the subjects in Group C were slight and quite unexceptional. Following training, correct responses and responses erring by one or two semitones occurred slightly more often, and responses erring three, four, and five semitones occurred slightly less often. Tritone errors, though rising slightly, appear relatively constant.

The effect of training on the individual subjects in Group C can be seen in Figure 25. Individual pretest and posttest performances are given for the number of correct responses, the mean absolute error, and the amount of information transmitted.
Figure 25. The effect of training on the individual scores of the number of correct responses, the mean absolute error, and the amount of information transmitted for the subjects in Group C.
The majority of subjects in Group C derived little benefit from their training efforts. Overall, the accuracy of Subjects B, I, and Z increased slightly, while the accuracy of Subjects E, G, L, and V decreased slightly. Considerable improvement, however, can be seen in the progress of Subjects F and K. One possible explanation of this phenomenon might be statistical regression toward the mean—a tendency for those who score highest on the pretest to score lower on the posttest, and those who score lower on the pretest to score higher on the posttest. Interpretation of the success of subjects F and K is problematic because both were among the poorest scorers on the pretest and therefore had more room for improvement.

Inspection of the individual error analysis by pitch class for Subject F revealed considerable improvement on the chromata D, F, F#, G, G#, A, and B following training, especially D, F, and G. Perhaps this could be explained by the fact that the practice log of Subject F revealed that he progressed to Lesson 5 (the highest level attained by any member of Group C) by only the fourth week of the experiment (a feat matched by only 10 of the 26 subjects, including all the members of Group A). The substantial improvement of Subject K is more difficult to explain. Inspection of the individual error analysis by pitch class for Subject K revealed dramatic improvement on the chromata C, D, D#, E, F#, G#, A, and A# following training, especially D, E and F#. This improvement is puzzling in light of the fact that the practice log of Subject K revealed that she practiced a total of only 40 minutes during two sessions and advanced only to Lesson 2. No other subject practiced for so few sessions and for less total time, made less progress, or discontinued participation earlier.
Discussion of Results Obtained from Group B

Group B was composed of 12 subjects: Subjects A, C, D, H, M, N, O, P, Q, S, W and Y. Each trained more than 1,200 minutes, and in fact, none trained less than 1,308 minutes. They trained a combined total of 26,749 minutes, distributed over a total of 1,219 training sessions. The mean number of training sessions completed per subject was 101.6 (median = 76, range = 47 to 191), and the mean number of minutes of training per subject was 2,229.1 (median = 1,722.5, range = 1,308 to 4,479). These twelve subjects accounted for 46.2% of the experimental group population, and their efforts amounted to 53.6% of all practice sessions and 51.1% of all practice time. Nevertheless, none reached the training criterion (advancement to Lesson 15) despite their persistence, and seven of them were still continuing to train at the conclusion of the training period. The mean highest difficulty level attained by the subjects in Group B was Lesson 8. The data obtained from the measurement of the mean number of correct responses, the mean absolute error, and the amount of information transmitted for Group B and the control group are contained in Table 18.
Table 18. Mean number of correct responses (120 possible), mean absolute error (semitones), mean amount of information transmitted (bits), their standard deviations, and number of observations on the pretest and the posttest for Group B and the control group.

<table>
<thead>
<tr>
<th></th>
<th>Group B</th>
<th>Control Group</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pretest</td>
<td>Posttest</td>
<td>Pretest</td>
<td>Posttest</td>
</tr>
<tr>
<td>Number Correct</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>11.917</td>
<td>18.000</td>
<td>12.583</td>
<td>12.750</td>
</tr>
<tr>
<td>SD</td>
<td>5.035</td>
<td>9.686</td>
<td>8.150</td>
<td>5.236</td>
</tr>
<tr>
<td>N</td>
<td>12</td>
<td>12</td>
<td>24</td>
<td>24</td>
</tr>
<tr>
<td>Mean Absolute Error</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>2.949</td>
<td>2.672</td>
<td>2.875</td>
<td>2.907</td>
</tr>
<tr>
<td>SD</td>
<td>0.362</td>
<td>0.475</td>
<td>0.331</td>
<td>0.262</td>
</tr>
<tr>
<td>N</td>
<td>12</td>
<td>12</td>
<td>24</td>
<td>24</td>
</tr>
<tr>
<td>Amount of Information Transmitted</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>1.098</td>
<td>1.287</td>
<td>1.140</td>
<td>1.129</td>
</tr>
<tr>
<td>SD</td>
<td>0.212</td>
<td>0.305</td>
<td>0.224</td>
<td>0.162</td>
</tr>
<tr>
<td>N</td>
<td>12</td>
<td>12</td>
<td>24</td>
<td>24</td>
</tr>
</tbody>
</table>

To help convey the relationship between the two groups, the group means contained in Table 18 are shown in Figure 26.
Figure 26. Mean number correct, mean absolute error, and mean amount of information transmitted for subjects in Group B and the control group.
Student’s t tests for independent samples were used to determine whether the differences among posttest group means were statistically significant at the .05 probability level. The posttest means of the subjects in Group B indicate an improvement as a result of training that was significant for the total number correct \((t = 2.124, df = 34, p = .041)\) and for the mean amount of information transmitted \((t = 2.040, df = 34, p = .049)\), but not for the mean absolute error \((t = -1.926, df = 34, p = .063)\). Perhaps the deficit that these subjects incurred on the pretest was too great to overcome with the current probability level. However, one could still conclude (with a 93.7% certainty) that, following training, the difference in the absolute error means of Group B and of the control group was not the product of chance.

The substantial improvement displayed by Group B subjects from their pretest to their posttest was separated according to pitch class and analyzed. Group means for the 12 subjects were tabulated and plotted as a function of pitch class, and the results are shown in Figure 27.
Figure 27. Mean number of correct responses, mean absolute error, and mean amount of information transmitted as a function of pitch class for the subjects in Group B.
Figure 27 indicates that the responses of Group B on the posttest were noticeably improved when compared to those on the pretest. Clearly, seven of the twelve possible chroma were more accurately judged on the posttest, and the four that were the most accurately judged were A, F#, C, and E♭, respectively. These are the first four chroma undertaken in training, and therefore received the greatest amount of practice. No striking improvement was found on the chroma D, as had been seen in the results for Group C.

All pretest and posttest responses for the subjects in Group B were separated according to the size of their deviation from the correct response in semitones and then analyzed. Group means were tabulated and plotted as a function of error interval and absolute error interval, and the results are shown in Figures 28 and 29.
Figure 28. Mean number of responses as a function of error interval for the subjects in Group B.

Figure 29. Mean number of responses as a function of absolute error interval for the subjects in Group B.
The frequency distributions in Figures 28 and 29 indicate that, from the pretest to the posttest, the changes in the mean number of responses as a function of error interval and absolute error interval for the subjects in Group B were more pronounced than those of Group C. Following training, numbers of correct responses and two-semitone errors rose substantially, and the number of responses erring four and five semitones fell substantially. The number of responses erring by one and three semitones occurred slightly less often, while, as seen before, the number of tritone errors remained relatively constant.

The effect of training on the individual subjects in Group B can be seen in Figure 30. Individual pretest and posttest performances are given for the number of correct responses, the mean absolute error, and the amount of information transmitted.
Figure 30. The effect of training on the individual scores of the number of correct responses, the mean absolute error, and the amount of information transmitted for the subjects in Group B.
Figure 30 shows that two of the subjects in Group B—Subjects N and C—derived great benefits as a result of their training efforts. Overall, the accuracy of Subjects A, D, H, O, Q, S, W, and Y increased slightly, while the accuracy of Subjects M and P decreased slightly. The two cases of dramatic improvement, eight of slight improvement and two of slight decrease in ability found in Group B contrast strikingly with the two cases of dramatic improvement, three of slight improvement and four of slight decrease in ability reported for Group C. The performances of Subjects C and N clearly cannot be explained, in this case, in terms of statistical regression toward the mean, because neither of these subjects were among the poorest scorers on the pretest.

Inspection of the individual error analysis by pitch class for Subject N revealed dramatic improvement for the chromata A, F#, C, and Eb following training. On the posttest, Subject N correctly identified the pitch class A 10 out of 10 times, and the pitch class F# 9 out of 10 times. In fact, substantial improvement occurred not only on every chroma he had studied, but on several not yet undertaken. Improvement was noted on every chroma except F and B, but this is not altogether surprising because these two would not have been undertaken until Lessons 12 and 14, respectively, and Subject N advanced only to Lesson 10.

In comparison, the improvement of Subject C mirrored that of Subject N, but the effect was somewhat less pronounced. Inspection of the individual error analysis by pitch class for Subject C revealed substantial improvement primarily for the chromata A, C, Eb, and F#. Improvement occurred on every chroma except C#, D, and B, but this result can not be explained quite as easily
as before, because these chromata would have been undertaken in Lessons 7, 11 and 14, respectively, and Subject C did advance to Lesson 13.
Discussion of Results Obtained from Group A

Group A was composed of 5 subjects in total: Subjects J, R, T, U and X. Each reached the training criterion (Lesson 15) by the conclusion of the training period, the goal to which all members of the experimental group had pledged themselves before the experiment began. They trained a combined total of 21,135 minutes, distributed over a total of 819 training sessions. The mean number of training sessions completed per subject was 163.8 (median = 137, range = 91 to 230), and the mean number of minutes of training per subject was 4,227 (median = 3,408, range = 2,009 to 6,600). While these five subjects accounted for only 19.2% of the experimental group population, their efforts amounted to 36.0% of all practice sessions and 40.3% of all practice time. The data obtained from the measurement of the mean number of correct responses, the mean absolute error, and the amount of information transmitted for Group A and the control group are contained in Table 19.
Table 19. Mean number of correct responses (120 possible), mean absolute error (semitones), mean amount of information transmitted (bits), their standard deviations, and number of observations on the pretest and the posttest for Group A and the control group.

<table>
<thead>
<tr>
<th></th>
<th>Group A</th>
<th>Control Group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pretest</td>
<td>Posttest</td>
</tr>
<tr>
<td>Number Correct</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>30.800</td>
<td>59.000</td>
</tr>
<tr>
<td>SD</td>
<td>19.110</td>
<td>31.591</td>
</tr>
<tr>
<td>N</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Mean Absolute Error</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>2.028</td>
<td>1.125</td>
</tr>
<tr>
<td>SD</td>
<td>0.584</td>
<td>0.768</td>
</tr>
<tr>
<td>N</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Amount of Information Transmitted</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>1.713</td>
<td>2.431</td>
</tr>
<tr>
<td>SD</td>
<td>0.448</td>
<td>0.680</td>
</tr>
<tr>
<td>N</td>
<td>5</td>
<td>5</td>
</tr>
</tbody>
</table>

To help convey the relationship between the two groups, the group means contained in Table 19 are shown in Figure 31.
Figure 31. Mean number correct, mean absolute error, and mean amount of information transmitted for subjects in Group A and the control group.
Student's t tests for independent samples were used to determine whether the differences among posttest group means were statistically significant at the .05 probability level. Group A's posttest means clearly indicate dramatic improvement as a result of training. This improvement was found to be significant for all measures of absolute pitch ability; the total number correct ($t = 7.190, df = 27, p < .001$), the mean absolute error ($t = -9.488, df = 27, p < .001$), and the mean amount of information transmitted ($t = 8.783, df = 27, p < .001$).

The remarkable improvement displayed by Group A from their pretest to their posttest was separated according to pitch class and analyzed. Group means for the 5 subjects were tabulated and plotted as a function of pitch class. The results are shown in Figure 32.
Figure 32. Mean number of correct responses, mean absolute error, and mean amount of information transmitted as a function of pitch class for the subjects in Group A.
Figure 32 indicates that the posttest responses of the subjects in Group A were vastly improved compared to their pretest responses. Improvement was greater and more evenly distributed among the chromata than that of either Group C or B. All twelve chromata were judged more accurately, and the sharp peaks on A, F#, C, and Eb, found in Group B, were therefore lacking. Likewise, no evidence of extraordinary improvement on the chroma D, as displayed by Group C, was found.

All pretest and posttest responses for the subjects in Group A were separated according to the size of their deviation from the correct response in semitones and then analyzed. Group means were tabulated and plotted as a function of error interval and absolute error interval, and the results are shown in Figures 33 and 34.
Figure 33. Mean number of responses as a function of error interval for the subjects in Group A.

Figure 34. Mean number of responses as a function of absolute error interval for the subjects in Group A.
The frequency distributions in Figures 33 and 34 indicate that, from the pretest to the posttest, the changes in the mean number of responses as a function of error interval and absolute error interval for the subjects in Group A were much more pronounced than those of either Group C or Group B. Following training, numbers of correct responses rose dramatically, while the number of responses erring one semitone rose slightly. The number of responses erring by two, three, or five semitones occurred less frequently than before, and errors of four and six semitones (major third above and below and tritone) were almost completely eliminated.

The effect of training on the individual subjects in Group A can be seen in Figure 35. Individual pretest and posttest performances are given for the number of correct responses, the mean absolute error, and the amount of information transmitted.
Figure 35. The effect of training on the individual scores of the number of correct responses, the mean absolute error, and the amount of information transmitted for the subjects in Group A.
Figure 35 shows that four of the five subjects in Group A benefited immensely from their training efforts. Following training, the overall accuracy of Subjects J, R, T, and U increased, while the accuracy of Subject X actually decreased. The 4 cases of dramatic improvement, and 1 case of a decrease in ability found in Group A contrast strikingly with the changes in ability reported for Groups C and B.

Inspection of the individual error analysis by pitch class for Subject R revealed a substantial improvement on all 12 chromata following training, though none could be identified with more than 50% accuracy. Subject J also showed improvement on all 12 chromata following training, and could identify four chromata, C, E, F, and F# with 70% accuracy or better. Subject U showed improvement on all 12 chromata following training, and could identified six chromata, C#, D, E, F, and F# and A with 70% accuracy or better. Subject T showed improvement on 10 of 12 chromata following training, committing an error on F on his posttest where he had previously been flawless, and remained perfect in his judgments of F# (he therefore had no possibility of improvement here). Following training, he identified 10 of the 12 chromata, all but E and B, with 80% accuracy or better and flawlessly identified 6 of them (C, C#, F#, G, G#, and A).

On the 120-item posttest, Subjects J, R, T, U, and X made 60, 30, 106, 69, and 30 correct responses (50%, 25%, 88.3%, 57.5% and 25%), respectively. The amount of chroma information they transmitted was 2.501, 1.846, 3.341, 2.782 and 1.687 bits, respectively. Carroll (1975) measured the speed and accuracy of four subjects who had possessed absolute pitch since childhood, and that of P. T. Brady, who claims to have acquired the ability through self-training. In Carroll’s experiment, the subjects’ task was to strike, as rapidly and accurately
as possible, the key of a non-muted grand piano, attempting to match pitches produced on a tape-recording. Tapes contained series of randomized piano tones in sets of up to 64 consecutive semitones. In a discussion of the results, Carroll wrote:

Pitch may be a two-dimensional continuum containing both chroma and octave information. It is thus possible that the high estimates of channel capacity for AP [absolute pitch] subjects should be viewed as sums of capacity values for these two dimensions. First, note that the values of IO [octave information] do not differ much between AP and NAP [non-absolute pitch] subjects. . . . The case is different, however, for chroma information, . . . The fact remains, however, that all AP subjects were found to transmit well over 2.5 bits of chroma information, usually from 3.0 to 3.4 bits, under the 64-note condition. (pp. 28-29)

Extrapolating from Carroll’s report, the actual scores of his 5 subjects with absolute pitch fell within a range from approximately 2.55 bits to approximately 3.45 bits, and Brady’s score was approximately 2.95 bits. Carroll reported the 5 subjects’ mean as 3.031 bits (p. 54). He also reported that their mean proportion of perfectly correct responses was 71.0% (actual scores were 93, 80, 67, 67, and 48, (p. 39)), and their mean proportion of responses with zero chroma error was 87.7% (p. 53). Subjects R and X did not appear to meet this standard. Perhaps Subjects J and U did, because their identification tests, while comprising only 48 notes instead of 64 notes, were completely devoid of feedback. Clearly, the posttest abilities of Subject T rivaled the performance of Carroll’s subjects, and surpassed that of Brady.

The performance of Subject X on the posttest was, without a doubt, more irregular than that of any other subject involved in the experiment. Inspection of her individual error analysis by pitch class revealed that of the 12 chromata, substantial improvement occurred on only one chroma (F#),
slight improvement occurred on 4 chromata, moderate declines in accuracy occurred on 3 chromata, and quite large declines in accuracy occurred on 4 chromata. On the posttest, only 2 chromata were identified with 70% accuracy or better, F# and A. What makes all of this even more puzzling was that inspection of the practice logs revealed that Subject X practiced more than any other subject in the experiment, measured in either total number of practice sessions or total duration of practice time.\textsuperscript{11}

\textsuperscript{11}Seeking confirmation that she had done well on the posttest, Subject X was disappointed in her score. She then mentioned that she had suffered from an ear infection a few weeks previously but, being completely symptom-free, believed herself to be fully recovered at the time she was administered the posttest. The experimenter reminded the subject that her posttest performance was not necessarily a valid indication of her musical abilities, and that all information concerning her identity would remain in strict confidence.
The Correlational Method

Several interesting relationships were found to exist by assessing the strength of the association between certain experimental group variables. What was the relationship between absolute pitch ability after training and the amount of training received? Coefficients of correlation using the Pearson $r$ were tabulated using the posttest scores, measured in terms of amount of information transmitted in bits, for the experimental group. Moderate-strength correlations were found upon comparison of the posttest results to the total numbers of practice sessions ($r = .423; t = 2.285, df = 25, p = .031$), and the total duration of practice time in minutes ($r = .466; t = 2.578, df = 25, p = .016$). A high-strength correlation was found upon comparison of the posttest results to the highest level of progress attained in the practice sessions ($r = .728; t = 5.208, df = 25, p < .001$).

What was the relationship between the variables affecting the level of progress attained in the training method? Coefficients of correlation using the Pearson $r$ were tabulated using the scores of the pretest for the experimental group. Correlations of moderate strengths were found upon comparison of the highest level of progress attained in the training method and the pretest numbers of correct responses ($r = .483; t = 2.705, df = 25, p = .012$), the pretest mean absolute error in semitones ($r = -.523; t = -3.006, df = 25, p = .006$), and the pretest amount of information transmitted in bits ($r = .531; t = 3.068, df = 25, p = .005$). High-strength correlations were found upon comparison of the highest level of progress attained in the training method to the total numbers of practice sessions ($r = .705; t = 4.875, df = 25, p < .001$), and the total duration of practice time in minutes ($r = .711; t = 4.958, df = 25, p < .001$).
Lastly, what was the relationship between absolute pitch ability after training and native absolute pitch ability, measured before training began? No predictions had been made about this relationship. Coefficients of correlation using the Pearson $r$ were tabulated using the scores of the pretests and posttests for the experimental group. Correlations of relatively high strengths were found upon comparison of the pretest and posttest numbers of correct responses ($r = .746; t = 5.486, df = 25, p < .001$), the pretest and posttest mean absolute error in semitones ($r = .746; t = 5.491, df = 25, p < .001$), and the pretest and posttest amount of information transmitted in bits ($r = .773; t = 5.962, df = 25, p < .001$).

To interpret these coefficients of correlation correctly, it is important to realize that the size of $r$ does not directly reflect the percentage of the association. For example, an $r = .700$ is not in any way "twice as strong a relationship" indicated by $r = .350$. Likewise, the difference between a correlation coefficient of $r = .400$ and $r = .500$ is not the same as the difference between a correlation coefficient of $r = .800$ and $r = .900$. Perhaps a more common misinterpretation of coefficients of correlation is to infer causation. Coefficients of correlation indicate that the variation of one variable is associated in some manner with the variation of some other, it does not imply that one causes the other.
CHAPTER VI

SUMMARY

Internal and External Validity

All experimenters face the same dilemma concerning the internal and external validity of their studies. Both factors are important, and often an attempt to increase one results in a corresponding decrease in the other. The experimental methodology detailed in Chapter IV was the result of a dedicated effort to minimize threats to the internal validity of the experiment. Without proper control, the effects of extraneous variables could confound the effect of the independent variable on the dependent variable; the resulting data could have been uninterpretable. This was not the case, and the data presented in Chapter V support the finding that the changes in the abilities of the subjects in the experimental group are the result of training. The matter of external validity raises the question of the generalizability of the findings of a study to other populations, in different settings, who would be affected by the variables that accompany non-identical treatments and measurements. This issue is a more subjective matter, and is often more difficult to resolve.
Generalizability of the Results of the Study

In the narrowest possible scope, the results of this experiment could be safely generalized to other groups of pretested, recruited, upper-level undergraduate music majors with two years of aural training, at Midwestern universities with similar academic and music performance/entrance standards, who have volunteered their participation in exchange for similar monetary rewards, if they are trained and tested in manners similar to that found in this experiment. To expand this experiment's findings beyond this limited compass, three factors should be considered when multiple-treatment interference is not a consideration. These include: (a) the nonreactive nature of the pretest; i.e., the administration of an 120-item pitch identification pretest (without correction) would not make the results obtained from pretested subjects unrepresentative from those of unpretested subjects; (b) the interaction effects of selection bias and the independent variable; i.e., not thought to be a significant factor since most potential trainees would be volunteers interested in acquiring absolute pitch; and (c) the reactive effects of the experimental setting; i.e., thought to be minimal in this study due to the latitude afforded the subjects regarding the times and places of their training sessions, and the non-threatening nature of the testing sessions. The generalizability of the results of this experiment are therefore seen as extending far beyond the worst-case scenario discussed earlier. In the light of these comments, the following conclusions are offered.

\[1\text{Campbell & Stanley, 1963, pp. 5-6.}\]
Conclusions

The results of the training experiment indicate that all of the hypotheses set out in Chapter III were supported. Specifically:

1. Nonpossessors of absolute pitch lack chroma identification skills. The pretest means of the experimental and control groups, though better than random, were unimpressive. Prior to training, the differences between the two groups were not statistically significant for any measure of absolute pitch ability: i.e., the total number correct, the mean absolute error, or the mean amount of information transmitted.

2. Absolute pitch identification skills do not increase, at least among adults, without training. The posttest means of the control group were remarkably close to their pretest means. The effects of maturation, in terms of completing another academic year as a college music major, had no positive effect on pitch identification abilities within the control group.

3. Absolute pitch identification skills can be improved with training. The posttest means of the experimental group were superior to those of the control group. The differences between groups were found to be statistically significant for all measures of absolute pitch ability: i.e., the total number correct, the mean absolute error, and the mean amount of information transmitted.
4. Advancement in the training method was related to the amount of effort expended by the subjects; strong correlations were found. Correlations between advancement in the training method and measures of pretest ability were of only moderate strength.

5. Absolute pitch ability following training was related to advancement in the training method; a strong correlation was found. Correlations between the posttest scores and measures of the amount of effort expended by the subjects were only moderately strong. It was also found that correlations between absolute pitch ability and native ability were slightly stronger than any of the other correlations, though this finding had not been anticipated.

Critics will undoubtedly point out that the posttest means of the complete experimental group were substantially below the levels of performance normally associated with native possessors of absolute pitch. Acknowledging this fact, the author offers the explanation that the complete experimental group means contain data derived from subjects with vastly differing levels of participation. Nonetheless, the results obtained from the entire experimental and control groups lend convincing support to the position that absolute pitch skills can be improved by practice. Moreover, results obtained after a division of the experimental group into three smaller groups according to their respective levels of participation lends support to the position that increased levels of the treatment, the training, were the cause of increased levels of posttest performance; this conclusion is also
supported by the correlational data. The correlational data suggest strong associations among certain variables; they suggest, but do not establish, causation.

Three observations may be made:

1. Posttest performance is more closely related to progress in the training method than to the amount of training effort expended by the subjects.

2. Progress in the training method is more closely related to the amount of training effort expended by the subjects than to their pretest performance.

3. Posttest scores were, nonetheless, as strongly related to the pretest scores as they were to progress in the training method.

One interpretation of these observations seems most reasonable. Many of the subjects who attained high scores on the posttest were the ones who completed the training regimen; however, different subjects progressed at different rates. Some subjects devoted substantial effort to their training, yet made little progress. Conversely, other subjects made extraordinary amounts of progress relative to the amount of effort they expended at training. Obviously, those who discontinued their participation early in the experiment stood little chance of completing the training regimen. Regardless, posttest performance is as strongly associated with pretest performance as with progress in the training method. The range of scores on the pretest provide evidence that there is a continuum of ability levels, rather than a bi-modal split, between nonpossessors and possessors of
absolute pitch. There appeared to be a strong tendency among those with better scores on the pretest to progress more quickly and more easily through the training program.

Finally, the question that many readers will ask is whether, in the course of this experiment, anyone actually acquired absolute pitch. Results obtained from the subjects who completed the training regimen, Group A, indicate that at least one and possibly three of the five subjects who completed the training regimen actually became possessors. Subjects J, T, and U made 60, 106, and 69 correct responses on the 120-item posttest without feedback (50%, 88.3%, and 57.5%), and 90, 112, and 105 of their responses were of one semitone error or less (75%, 93.3%, and 87.5%), respectively. The amount of chroma information they transmitted was 2.501, 3.341, and 2.782 bits (of a maximum 3.585 bits), respectively. The performances of Subjects J and U approach the level of Carroll’s (1975) absolute pitch subjects, even though Carroll’s subjects received feedback. Despite the obvious disadvantage of not having feedback, the performance of Subject T rivals Carroll’s best performers and surpasses the performance of Brady, whose 1970 self-training experiment is probably the most often cited example of successful training in the literature. The other two subjects in Group A did not attain this level of performance.

---

2Interestingly, Subject J made a single tritone error; Subjects T and U made none.
Implications

The research problem of attempting to instill absolute pitch skills through training reaches to the very heart of aural training in this country and abroad. The two major schools of thought regarding sight-singing instruction (the fixed-do and movable-do systems of solfège) differ fundamentally on the basis of attention paid to the absolute identities of musical pitches. Of course, the underlying implication of this experiment is that absolute pitch can indeed be learned. As detailed in Chapter II, close inspection of almost any previous training experiment will reveal a serious flaw either in its design, or in its execution, or both. Nonetheless, most of these studies report that, following training, test subjects improved their absolute pitch abilities. This experiment is offered as the first long-term training study of rigorous experimental design, with a sufficiently large sample of musical subjects training with musical stimuli while receiving feedback, measured with reliable and valid tests and conducted by an experimenter who is not a subject, who does not have absolute pitch, and therefore has nothing to prove or disprove about the ability. This study lends strong support to those who advocate the role of chroma in judgments of absolute pitch, and offers convincing evidence for the learning theory of absolute pitch; the odds of these experimental results occurring by chance alone are statistically insignificant.

The point that remains to be made is that—assuming absolute pitch can be taught—never have we had a better opportunity than now to teach it. The advent of powerful personal computers, relatively inexpensive synthesizers and samplers, and user-friendly interfaces and programming languages should make college aural skills labs more effective today than
ever before. However, the decision as to whether the benefits of possessing absolute pitch are justified by the time required for its acquisition, is most likely a personal one, and has been purposely avoided in this investigation. This question, of course, is loaded; it implies that there are significant musical benefits to possessing absolute pitch, a highly debatable point in itself.

**Recommendations for Future Study**

Despite the effectiveness of the training method demonstrated in this experiment, serious reservations remain regarding the value of its adoption for college teaching. This author hesitates to recommend the handbook without extensive revision to eradicate a host of mistakes and inaccuracies, unsubstantiated beliefs and unsupported claims, and to bring its text into accordance with the accepted terminology of the field. By uncritical adoption or endorsement of this book in its present form, much of the care taken to teach students the advantages of scientific methodology, the rewards of scholarly rigor, and the power of precise writing could be lost.

Replication of this training experiment could help to confirm the effect of training reported here, as well as to establish the effect's generalizability. While inclusion of a control group usually reduces the available number of subjects—a statistical disadvantage—future experimenters are urged to include them. Furthermore, future experimenters are encouraged to standardize their tests of absolute pitch ability, as well as standardizing their method of scoring them. It is hoped that the effort to model the test used in this experiment after Ward’s absolute pitch test, and to score it in accordance with Carroll’s method of determining the amount of chroma information
transmitted, proved worthwhile in making the results of this experiment more generalizable to the work of others.

Finally, in an effort to determine whether there are significant musical benefits to possessing absolute pitch, future experimenters might consider measuring the musical abilities of a group of randomly chosen subjects, randomly dividing the group, and randomly assigning treatment (absolute pitch training). After the experimental group had acquired absolute pitch, a posttest would be given to reassess the musical abilities measured earlier. Perhaps only then could the mystical claims of the superiority of absolute pitch possessors be supported—or laid to rest.
BIBLIOGRAPHY


Abraham, O. (1907). Das absolute Tonbewußtsein und die Musik [Absolute pitch and music]. Sammelbände der Internationalen Musikgesellschaft, 8, 486-491.


Aizawa, M. (1959). Zettai onkō handon to onkō shikibetsusei oyobi chinō to no kankei ni tsuite [On the relations of pitch discrimination and intelligence to the judgment of absolute pitch]. [Psychological papers to commemorate the 35th anniversary of Dr. Yoshikazu Ohwaki’s professorship at Tōhoku University, Sendai, Japan], 74-82.


Auerbach, F. (1906). Das absolute Tonbewußtsein und die Musik [Absolute pitch and music]. *Sammelbände der Internationalen Musikgesellschaft, 8*, 105-112.


Cuddy, L. L. (1972). Comment on “Practice effects in the absolute judgment of frequency” by Heller and Auerbach. Psychonomic Science, 28, 68.


APPENDIX A

MATERIALS RELATIVE TO CHAPTER IV
Subject Personal Data Questionnaire

Name:
Campus Address:
Campus Phone:
Work Phone:
Social Security or Student I.D. number:

Permanent Street Address:
Permanent City and State:
Permanent Phone:

Academic Major:
Class Rank:
Principal Instrument (or voice range):
Age:
Gender:

Have you completed Music 426?
Will you still be attending Ohio State during Spring Quarter, 1988?
Do you think that you have absolute pitch?
At what age did you start your musical training?
On what instrument did you begin this training?
Which hand are you writing with right now?
Are you taking lessons on your instrument this year?
Are you a member of any type of performing ensemble this year?
If so, will this ensemble give a public performance this year?
Will you perform in any type of solo performance this year?
Are you involved with musical activities on a daily basis?
Pitch Identification Test

Block One

\[ \text{Duration: } d = 60 \]
Pitch Identification Test

Block Two
Pitch Identification Test

Block Three
Pitch Identification Test

Block Four
Pitch Identification Test

Block Five
Pitch Identification Test Directions

You are about to take a test designed to measure absolute pitch naming ability in adult musicians. You will be presented five blocks of 24 test items, for a total of 120 items. Each individual item will consist of a two-second tone followed by a maximum of three seconds in which to respond. There will be a 20-second rest period between each block.

Once a test tone is heard, write down at once its pitch name just as you feel it to be. Do not attempt to remember a tone after it has been judged, and judge each tone independently, without reference to any other tones. Just give your first, automatic, intuitive, "gut-hunch" pitch name; do not work hard at the task, and do not use relative pitch.

Do not hum, sing, whistle or talk once the test has begun. You must give an answer for each of the 120 test items, and do not attempt to change or check your answers during the course of the test or after the test has been completed.

Are there any questions? Let's begin.
Subject Consent Form For Participation
In Music Perception Research

THE OHIO STATE UNIVERSITY

Protocol No. 76B0486

I consent to participate as a subject in research entitled: AN EXPERIMENTAL INVESTIGATION OF THE EFFECTIVENESS OF TRAINING ON THE ACQUISITION OF ABSOLUTE PITCH IN ADULT MUSICIANS.

Mark Rush has explained the purpose of the study, the procedures to be followed, and the expected duration of my participation. Possible benefits of the study have been described as have alternative procedures, if such procedures are applicable and available. Possible risks or complications which may be reasonably expected have been explained to me, and I understand them.

While I am a subject in this experiment I agree to faithfully commit 25 minutes a day, everyday, to practicing the experimental techniques. I agree to practice at times when I am mentally alert, and agree to abstain from the use of recreational drugs throughout the duration of the study.

I acknowledge that I have had the opportunity to obtain additional information regarding the study and that any questions I have raised have been answered to my full satisfaction. Further, I understand that I am free to withdraw my consent at any time and to discontinue participation in the study without fear of prejudice or harassment. I understand that all information obtained from me concerning my identity and concerning my performance on test instruments will remain in strict confidence. I also understand that my performance on test instruments is not necessarily a valid indication of my musical abilities.

Finally, I acknowledge that I have read and fully understand the consent form. I sign it freely and voluntarily and understand a copy is available to me at my request.

Date: __________ Signed: ___________________________

Time: ________ (AM/PM) Witness: __________________________

Signed: __________________________
(Principal Investigator)

HS-027 (Rev. 10-87)—To be used only in connection with social and behavioral research.
Absolute Pitch
Daily Practice Log

Name: 
Partner's Name: 

Sun., Feb. 21, 1988 
Start: (am/pm) 
End: (am/pm) 
Exercise Number: 
Total Practice Session Length in Minutes: 
Today's Practice Instrument: 
Practice Location: 
This practice session was a (Team/Solo) effort.

Mon., Feb. 22, 1988 
Start: (am/pm) 
End: (am/pm) 
Exercise Number: 
Total Practice Session Length in Minutes: 
Today's Practice Instrument: 
Practice Location: 
This practice session was a (Team/Solo) effort.

Tues, Feb. 23, 1988 
Start: (am/pm) 
End: (am/pm) 
Exercise Number: 
Total Practice Session Length in Minutes: 
Today's Practice Instrument: 
Practice Location: 
This practice session was a (Team/Solo) effort.

Wed., Feb. 24, 1988 
Start: (am/pm) 
End: (am/pm) 
Exercise Number: 
Total Practice Session Length in Minutes: 
Today's Practice Instrument: 
Practice Location: 
This practice session was a (Team/Solo) effort.

Thurs., Feb. 25, 1988 
Start: (am/pm) 
End: (am/pm) 
Exercise Number: 
Total Practice Session Length in Minutes: 
Today's Practice Instrument: 
Practice Location: 
This practice session was a (Team/Solo) effort.

Fri., Feb. 26, 1988 
Start: (am/pm) 
End: (am/pm) 
Exercise Number: 
Total Practice Session Length in Minutes: 
Today's Practice Instrument: 
Practice Location: 
This practice session was a (Team/Solo) effort.

Sat., Feb. 27, 1988 
Start: (am/pm) 
End: (am/pm) 
Exercise Number: 
Total Practice Session Length in Minutes: 
Today's Practice Instrument: 
Practice Location: 
This practice session was a (Team/Solo) effort.

The practice sessions listed above were conducted while I was mentally and physically alert, and free from the use of recreational drugs.

Signed: ____________________________
APPENDIX B

INDIVIDUAL ERROR ANALYSES BY PITCH CLASS
Figure 36. Number correct, mean absolute error, and information transmitted as a function of pitch class for Subject A in the experimental group.
Figure 37. Number correct, mean absolute error, and information transmitted as a function of pitch class for Subject B in the experimental group.
Figure 38. Number correct, mean absolute error, and information transmitted as a function of pitch class for Subject C in the experimental group.
Figure 39. Number correct, mean absolute error, and information transmitted as a function of pitch class for Subject D in the experimental group.
Figure 40. Number correct, mean absolute error, and information transmitted as a function of pitch class for Subject E in the experimental group.
Figure 41. Number correct, mean absolute error, and information transmitted as a function of pitch class for Subject F in the experimental group.
Figure 42. Number correct, mean absolute error, and information transmitted as a function of pitch class for Subject G in the experimental group.
Figure 43. Number correct, mean absolute error, and information transmitted as a function of pitch class for Subject H in the experimental group.
Figure 44. Number correct, mean absolute error, and information transmitted as a function of pitch class for Subject I in the experimental group.
Figure 45. Number correct, mean absolute error, and information transmitted as a function of pitch class for Subject J in the experimental group.
Figure 46. Number correct, mean absolute error, and information transmitted as a function of pitch class for Subject K in the experimental group.
Figure 47. Number correct, mean absolute error, and information transmitted as a function of pitch class for Subject L in the experimental group.
Figure 48. Number correct, mean absolute error, and information transmitted as a function of pitch class for Subject M in the experimental group.
Figure 49. Number correct, mean absolute error, and information transmitted as a function of pitch class for Subject N in the experimental group.
Figure 50. Number correct, mean absolute error, and information transmitted as a function of pitch class for Subject O in the experimental group.
Figure 51. Number correct, mean absolute error, and information transmitted as a function of pitch class for Subject P in the experimental group.
Figure 52. Number correct, mean absolute error, and information transmitted as a function of pitch class for Subject Q in the experimental group.
Figure 53. Number correct, mean absolute error, and information transmitted as a function of pitch class for Subject R in the experimental group.
Figure 54. Number correct, mean absolute error, and information transmitted as a function of pitch class for Subject S in the experimental group.
Figure 55. Number correct, mean absolute error, and information transmitted as a function of pitch class for Subject T in the experimental group.
Figure 56. Number correct, mean absolute error, and information transmitted as a function of pitch class for Subject U in the experimental group.
**Figure 57.** Number correct, mean absolute error, and information transmitted as a function of pitch class for Subject V in the experimental group.
Figure 58. Number correct, mean absolute error, and information transmitted as a function of pitch class for Subject W in the experimental group.
Figure 59. Number correct, mean absolute error, and information transmitted as a function of pitch class for Subject X in the experimental group.
Figure 60. Number correct, mean absolute error, and information transmitted as a function of pitch class for Subject Y in the experimental group.
Figure 61. Number correct, mean absolute error, and information transmitted as a function of pitch class for Subject Z in the experimental group.
Figure 62. Number correct, mean absolute error, and information transmitted as a function of pitch class for Subject AA in the control group.
Figure 63. Number correct, mean absolute error, and information transmitted as a function of pitch class for Subject BB in the control group.
Figure 64. Number correct, mean absolute error, and information transmitted as a function of pitch class for Subject CC in the control group.
Figure 65. Number correct, mean absolute error, and information transmitted as a function of pitch class for Subject DD in the control group.
Figure 66. Number correct, mean absolute error, and information transmitted as a function of pitch class for Subject EE in the control group.
Figure 67. Number correct, mean absolute error, and information transmitted as a function of pitch class for Subject FF in the control group.
Figure 68. Number correct, mean absolute error, and information transmitted as a function of pitch class for Subject GG in the control group.
Figure 69. Number correct, mean absolute error, and information transmitted as a function of pitch class for Subject HH in the control group.
Figure 70. Number correct, mean absolute error, and information transmitted as a function of pitch class for Subject II in the control group.
Figure 71. Number correct, mean absolute error, and information transmitted as a function of pitch class for Subject JJ in the control group.
Figure 72. Number correct, mean absolute error, and information transmitted as a function of pitch class for Subject KK in the control group.
Figure 73. Number correct, mean absolute error, and information transmitted as a function of pitch class for Subject LL in the control group.
Figure 74. Number correct, mean absolute error, and information transmitted as a function of pitch class for Subject MM in the control group.
Figure 75. Number correct, mean absolute error, and information transmitted as a function of pitch class for Subject NN in the control group.
Figure 76. Number correct, mean absolute error, and information transmitted as a function of pitch class for Subject OO in the control group.
Figure 77. Number correct, mean absolute error, and information transmitted as a function of pitch class for Subject PP in the control group.
Figure 78. Number correct, mean absolute error, and information transmitted as a function of pitch class for Subject QQ in the control group.
Figure 79. Number correct, mean absolute error, and information transmitted as a function of pitch class for Subject RR in the control group.
Figure 80. Number correct, mean absolute error, and information transmitted as a function of pitch class for Subject SS in the control group.
Figure 81. Number correct, mean absolute error, and information transmitted as a function of pitch class for Subject TT in the control group.
Figure 82. Number correct, mean absolute error, and information transmitted as a function of pitch class for Subject UU in the control group.
Figure 83. Number correct, mean absolute error, and information transmitted as a function of pitch class for Subject VV in the control group.
Figure 84. Number correct, mean absolute error, and information transmitted as a function of pitch class for Subject WW in the control group.
Figure 85. Number correct, mean absolute error, and information transmitted as a function of pitch class for Subject XX in the control group.
Figure 86. Number correct, mean absolute error, and information transmitted as a function of pitch class for Subject YY in the control group.
Figure 87. Number correct, mean absolute error, and information transmitted as a function of pitch class for Subject ZZ in the control group.
APPENDIX C

INDIVIDUAL ERROR ANALYSES BY INTERVAL CLASS
Figure 89. Number of responses as a function of error interval for Subject A in the experimental group.

Figure 89. Number of responses as a function of absolute error interval for Subject A in the experimental group.
Figure 90. Number of responses as a function of error interval for Subject B in the experimental group.

Figure 91. Number of responses as a function of absolute error interval for Subject B in the experimental group.
Figure 92. Number of responses as a function of error interval for Subject C in the experimental group.

Figure 93. Number of responses as a function of absolute error interval for Subject C in the experimental group.
**Figure 94.** Number of responses as a function of error interval for Subject D in the experimental group.

**Figure 95.** Number of responses as a function of absolute error interval for Subject D in the experimental group.
Figure 96. Number of responses as a function of error interval for Subject E in the experimental group.

Figure 97. Number of responses as a function of absolute error interval for Subject E in the experimental group.
Figure 98. Number of responses as a function of error interval for Subject F in the experimental group.

Figure 99. Number of responses as a function of absolute error interval for Subject F in the experimental group.
**Figure 100.** Number of responses as a function of error interval for Subject G in the experimental group.

**Figure 101.** Number of responses as a function of absolute error interval for Subject G in the experimental group.
Figure 102. Number of responses as a function of error interval for Subject H in the experimental group.

Figure 103. Number of responses as a function of absolute error interval for Subject H in the experimental group.
**Figure 104.** Number of responses as a function of error interval for Subject I in the experimental group.

**Figure 105.** Number of responses as a function of absolute error interval for Subject I in the experimental group.
Figure 106. Number of responses as a function of error interval for Subject J in the experimental group.

Figure 107. Number of responses as a function of absolute error interval for Subject J in the experimental group.
Figure 108. Number of responses as a function of error interval for Subject K in the experimental group.

Figure 109. Number of responses as a function of absolute error interval for Subject K in the experimental group.
Figure 110. Number of responses as a function of error interval for Subject L in the experimental group.

Figure 111. Number of responses as a function of absolute error interval for Subject L in the experimental group.
**Figure 112.** Number of responses as a function of error interval for Subject M in the experimental group.

**Figure 113.** Number of responses as a function of absolute error interval for Subject M in the experimental group.
Figure 114. Number of responses as a function of error interval for Subject N in the experimental group.

Figure 115. Number of responses as a function of absolute error interval for Subject N in the experimental group.
**Figure 116.** Number of responses as a function of error interval for Subject O in the experimental group.

**Figure 117.** Number of responses as a function of absolute error interval for Subject O in the experimental group.
Figure 118. Number of responses as a function of error interval for Subject P in the experimental group.

Figure 119. Number of responses as a function of absolute error interval for Subject P in the experimental group.
Figure 120. Number of responses as a function of error interval for Subject Q in the experimental group.

Figure 121. Number of responses as a function of absolute error interval for Subject Q in the experimental group.
**Figure 122.** Number of responses as a function of error interval for Subject R in the experimental group.

**Figure 123.** Number of responses as a function of absolute error interval for Subject R in the experimental group.
Figure 124. Number of responses as a function of error interval for Subject S in the experimental group.

Figure 125. Number of responses as a function of absolute error interval for Subject S in the experimental group.
Figure 126. Number of responses as a function of error interval for Subject T in the experimental group.

Figure 127. Number of responses as a function of absolute error interval for Subject T in the experimental group.
Figure 128. Number of responses as a function of error interval for Subject U in the experimental group.

Figure 129. Number of responses as a function of absolute error interval for Subject U in the experimental group.
**Figure 130.** Number of responses as a function of error interval for Subject V in the experimental group.

**Figure 131.** Number of responses as a function of absolute error interval for Subject V in the experimental group.
**Figure 132.** Number of responses as a function of error interval for Subject W in the experimental group.

**Figure 133.** Number of responses as a function of absolute error interval for Subject W in the experimental group.
Figure 134. Number of responses as a function of error interval for Subject X in the experimental group.

Figure 135. Number of responses as a function of absolute error interval for Subject X in the experimental group.
Figure 136. Number of responses as a function of error interval for Subject Y in the experimental group.

Figure 137. Number of responses as a function of absolute error interval for Subject Y in the experimental group.
**Figure 138.** Number of responses as a function of error interval for Subject Z in the experimental group.

**Figure 139.** Number of responses as a function of absolute error interval for Subject Z in the experimental group.
Figure 140. Number of responses as a function of error interval for Subject AA in the control group.

Figure 141. Number of responses as a function of absolute error interval for Subject AA in the control group.
**Figure 142.** Number of responses as a function of error interval for Subject BB in the control group.

**Figure 143.** Number of responses as a function of absolute error interval for Subject BB in the control group.
Figure 144. Number of responses as a function of error interval for Subject CC in the control group.

Figure 145. Number of responses as a function of absolute error interval for Subject CC in the control group.
Figure 146. Number of responses as a function of error interval for Subject DD in the control group.

Figure 147. Number of responses as a function of absolute error interval for Subject DD in the control group.
Figure 148. Number of responses as a function of error interval for Subject EE in the control group.

Figure 149. Number of responses as a function of absolute error interval for Subject EE in the control group.
Figure 150. Number of responses as a function of error interval for Subject FF in the control group.

Figure 151. Number of responses as a function of absolute error interval for Subject FF in the control group.
Figure 152. Number of responses as a function of error interval for Subject GG in the control group.

Figure 153. Number of responses as a function of absolute error interval for Subject GG in the control group.
Figure 154. Number of responses as a function of error interval for Subject HH in the control group.

Figure 155. Number of responses as a function of absolute error interval for Subject HH in the control group.
Figure 156. Number of responses as a function of error interval for Subject II in the control group.

Figure 157. Number of responses as a function of absolute error interval for Subject II in the control group.
Figure 158. Number of responses as a function of error interval for Subject JJ in the control group.

Figure 159. Number of responses as a function of absolute error interval for Subject JJ in the control group.
Figure 160. Number of responses as a function of error interval for Subject KK in the control group.

Figure 161. Number of responses as a function of absolute error interval for Subject KK in the control group.
Figure 162. Number of responses as a function of error interval for Subject LL in the control group.

Figure 163. Number of responses as a function of absolute error interval for Subject LL in the control group.
Figure 164. Number of responses as a function of error interval for Subject MM in the control group.

Figure 165. Number of responses as a function of absolute error interval for Subject MM in the control group.
Figure 166. Number of responses as a function of error interval for Subject NN in the control group.

Figure 167. Number of responses as a function of absolute error interval for Subject NN in the control group.
Figure 168. Number of responses as a function of error interval for Subject OO in the control group.

Figure 169. Number of responses as a function of absolute error interval for Subject OO in the control group.
Figure 170. Number of responses as a function of error interval for Subject PP in the control group.

Figure 171. Number of responses as a function of absolute error interval for Subject PP in the control group.
Figure 172. Number of responses as a function of error interval for Subject QQ in the control group.

Figure 173. Number of responses as a function of absolute error interval for Subject QQ in the control group.
Figure 174. Number of responses as a function of error interval for Subject RR in the control group.

Figure 175. Number of responses as a function of absolute error interval for Subject RR in the control group.
Figure 176. Number of responses as a function of error interval for Subject SS in the control group.

Figure 177. Number of responses as a function of absolute error interval for Subject SS in the control group.
Figure 178. Number of responses as a function of error interval for Subject TT in the control group.

Figure 179. Number of responses as a function of absolute error interval for Subject TT in the control group.
**Figure 180.** Number of responses as a function of error interval for Subject UU in the control group.

**Figure 181.** Number of responses as a function of absolute error interval for Subject UU in the control group.
Figure 182. Number of responses as a function of error interval for Subject VV in the control group.

Figure 183. Number of responses as a function of absolute error interval for Subject VV in the control group.
Figure 184. Number of responses as a function of error interval for Subject WW in the control group.

Figure 185. Number of responses as a function of absolute error interval for Subject WW in the control group.
Figure 186. Number of responses as a function of error interval for Subject XX in the control group.

Figure 187. Number of responses as a function of absolute error interval for Subject XX in the control group.
**Figure 188.** Number of responses as a function of error interval for Subject YY in the control group.

**Figure 189.** Number of responses as a function of absolute error interval for Subject YY in the control group.
Figure 190. Number of responses as a function of error interval for Subject ZZ in the control group.

Figure 191. Number of responses as a function of absolute error interval for Subject ZZ in the control group.
APPENDIX D

TRAINING LOG SUMMARIES
Figure 192. Total practice minutes, total practice sessions, and highest difficulty level practiced as a function of each week of the training period for Subject A in the experimental group.
Figure 193. Total practice minutes, total practice sessions, and highest difficulty level practiced as a function of each week of the training period for Subject B in the experimental group.
Figure 194. Total practice minutes, total practice sessions, and highest difficulty level practiced as a function of each week of the training period for Subject C in the experimental group.
Figure 195. Total practice minutes, total practice sessions, and highest difficulty level practiced as a function of each week of the training period for Subject D in the experimental group.
Figure 196. Total practice minutes, total practice sessions, and highest difficulty level practiced as a function of each week of the training period for Subject E in the experimental group.
Figure 197. Total practice minutes, total practice sessions, and highest difficulty level practiced as a function of each week of the training period for Subject F in the experimental group.
Figure 198. Total practice minutes, total practice sessions, and highest difficulty level practiced as a function of each week of the training period for Subject G in the experimental group.
Figure 199. Total practice minutes, total practice sessions, and highest difficulty level practiced as a function of each week of the training period for Subject H in the experimental group.
Figure 200. Total practice minutes, total practice sessions, and highest difficulty level practiced as a function of each week of the training period for Subject I in the experimental group.
Figure 201. Total practice minutes, total practice sessions, and highest difficulty level practiced as a function of each week of the training period for Subject J in the experimental group.
Figure 202. Total practice minutes, total practice sessions, and highest difficulty level practiced as a function of each week of the training period for Subject K in the experimental group.
Figure 203. Total practice minutes, total practice sessions, and highest difficulty level practiced as a function of each week of the training period for Subject L in the experimental group.
Figure 204. Total practice minutes, total practice sessions, and highest difficulty level practiced as a function of each week of the training period for Subject M in the experimental group.
Figure 205. Total practice minutes, total practice sessions, and highest difficulty level practiced as a function of each week of the training period for Subject N in the experimental group.
Figure 206. Total practice minutes, total practice sessions, and highest difficulty level practiced as a function of each week of the training period for Subject 0 in the experimental group.
Figure 207. Total practice minutes, total practice sessions, and highest difficulty level practiced as a function of each week of the training period for Subject P in the experimental group.
Figure 208. Total practice minutes, total practice sessions, and highest difficulty level practiced as a function of each week of the training period for Subject Q in the experimental group.
Figure 209. Total practice minutes, total practice sessions, and highest difficulty level practiced as a function of each week of the training period for Subject R in the experimental group.
Figure 210. Total practice minutes, total practice sessions, and highest difficulty level practiced as a function of each week of the training period for Subject S in the experimental group.
Figure 211. Total practice minutes, total practice sessions, and highest difficulty level practiced as a function of each week of the training period for Subject T in the experimental group.
**Figure 212.** Total practice minutes, total practice sessions, and highest difficulty level practiced as a function of each week of the training period for Subject U in the experimental group.
Figure 213. Total practice minutes, total practice sessions, and highest difficulty level practiced as a function of each week of the training period for Subject V in the experimental group.
Figure 214. Total practice minutes, total practice sessions, and highest difficulty level practiced as a function of each week of the training period for Subject W in the experimental group.
Figure 215. Total practice minutes, total practice sessions, and highest difficulty level practiced as a function of each week of the training period for Subject X in the experimental group.
Figure 216. Total practice minutes, total practice sessions, and highest difficulty level practiced as a function of each week of the training period for Subject Y in the experimental group.
Figure 217. Total practice minutes, total practice sessions, and highest difficulty level practiced as a function of each week of the training period for Subject Z in the experimental group.
Table 20. Summary of the total number of training sessions, their combined duration, the highest difficulty level practiced during the final week, and the week of the experiment in which it occurred for the subjects in the experimental group.

<table>
<thead>
<tr>
<th>Subject</th>
<th>Total Sessions (Number)</th>
<th>Total Duration (Minutes)</th>
<th>Final Level (Ex. Number)</th>
<th>Final Practice (Week)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subject A</td>
<td>171</td>
<td>3,836</td>
<td>6</td>
<td>34</td>
</tr>
<tr>
<td>Subject B</td>
<td>43</td>
<td>948</td>
<td>4</td>
<td>9</td>
</tr>
<tr>
<td>Subject C</td>
<td>92</td>
<td>1,997</td>
<td>13</td>
<td>34</td>
</tr>
<tr>
<td>Subject D</td>
<td>67</td>
<td>1,634</td>
<td>5</td>
<td>34</td>
</tr>
<tr>
<td>Subject E</td>
<td>36</td>
<td>483</td>
<td>4</td>
<td>14</td>
</tr>
<tr>
<td>Subject F</td>
<td>34</td>
<td>755</td>
<td>5</td>
<td>8</td>
</tr>
<tr>
<td>Subject G</td>
<td>38</td>
<td>322</td>
<td>4</td>
<td>16</td>
</tr>
<tr>
<td>Subject H</td>
<td>47</td>
<td>1,308</td>
<td>5</td>
<td>26</td>
</tr>
<tr>
<td>Subject I</td>
<td>41</td>
<td>930</td>
<td>4</td>
<td>13</td>
</tr>
<tr>
<td>Subject J</td>
<td>137</td>
<td>3,408</td>
<td>15</td>
<td>34</td>
</tr>
<tr>
<td>Subject K</td>
<td>2</td>
<td>40</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Subject L</td>
<td>9</td>
<td>210</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>Subject M</td>
<td>67</td>
<td>1,716</td>
<td>4</td>
<td>13</td>
</tr>
<tr>
<td>Subject N</td>
<td>127</td>
<td>2,835</td>
<td>10</td>
<td>34</td>
</tr>
<tr>
<td>Subject O</td>
<td>174</td>
<td>2,965</td>
<td>10</td>
<td>34</td>
</tr>
<tr>
<td>Subject P</td>
<td>191</td>
<td>4,479</td>
<td>6</td>
<td>34</td>
</tr>
<tr>
<td>Subject Q</td>
<td>77</td>
<td>1,330</td>
<td>5</td>
<td>13</td>
</tr>
<tr>
<td>Subject R</td>
<td>137</td>
<td>3,258</td>
<td>15</td>
<td>34</td>
</tr>
<tr>
<td>Subject S</td>
<td>73</td>
<td>1,729</td>
<td>5</td>
<td>34</td>
</tr>
<tr>
<td>Subject T</td>
<td>224</td>
<td>5,860</td>
<td>15</td>
<td>34</td>
</tr>
<tr>
<td>Subject U</td>
<td>91</td>
<td>2,009</td>
<td>15</td>
<td>34</td>
</tr>
<tr>
<td>Subject V</td>
<td>7</td>
<td>180</td>
<td>3</td>
<td>7</td>
</tr>
<tr>
<td>Subject W</td>
<td>75</td>
<td>1,555</td>
<td>10</td>
<td>28</td>
</tr>
<tr>
<td>Subject X</td>
<td>230</td>
<td>6,600</td>
<td>15</td>
<td>34</td>
</tr>
<tr>
<td>Subject Y</td>
<td>58</td>
<td>1,365</td>
<td>9</td>
<td>28</td>
</tr>
<tr>
<td>Subject Z</td>
<td>26</td>
<td>635</td>
<td>4</td>
<td>13</td>
</tr>
</tbody>
</table>
APPENDIX E

EXPERIMENTAL GROUP SUBJECTS’ COMMENTS
Subject A

My partner and I had difficulty in keeping the program going because of a lack of time. We remained quite consistent throughout the first three months, but starting in Winter Quarter (the fourth month), time became a problem. We had much difficulty in having our schedules coordinate, but we made it work as best as we could. The training came of benefit to me, the little that I did. I will pursue further training with this. I wish the whole experiment could have been a little more structured with a class-type meeting once a week or so. Oh well, thanks for the opportunity.

Subject B

I didn’t make it all the way through the program. About January it just started to become too much to get the practice time in with my recital approaching, and then my partner quit on me. Practicing alone just isn’t beneficial in my opinion. If I’d been able to continue, I do believe I could have progressed much further. It’s a good program, but if you don’t follow through with it, you keep having to start over at the beginning every time you go back to practicing.

Subject C

Although I have not progressed to the point that I can easily recognize pitches in a real musical situation, the study has forced me to concentrate more on using my ears as an evaluation tool. Thus, my hearing (relative pitch) has improved a marked degree since beginning the program. I plan on continuing to improve my pitch determining ability. Like learning a new language, it is something that I feel I can do, but will have to put forth a good deal of time to accomplish.

Subject D

As far as recognizing pitches in music, or sounds such as bells, I had no improvement of naming the pitches. But my overall ear developed; I heard things much clearer and sharper. Pitches were much more definite and focused. My relative pitch improved. It took a while for me to hear the colors that were associated with the pitches.

Subject E

At the beginning, I was really excited about this; I thought I could do it. But certain things came up that did not permit me to finish. I had a very busy class load the entire year and it was hard for my partner and me to practice. I wanted to make it work, but my partner didn’t seem as willing as I was. He frequently missed practice sessions we set up and the last three months of school I hardly saw him at all. I think I could have gone all the way through this if I had been able to practice every day.
Subject F

Not having finished the program, I have not acquired absolute pitch. However, I do feel that the time I spent made an impact. I hold a few pitches set in memory, but I have no musical use for this. In an isolated situation I can identify some notes, but out in the “real world” I could not really use it. The reason for my failure to continue in the program stems from timing schedules. My partner and I met regularly for all of Autumn Quarter but after that I could not get our schedules to meet. I never found a new partner, and thus discontinued the program.

Subject G

I dropped out because of partner conflict. He would only show up once in a while, then with schedule conflicts... it just didn’t work. With the short time involved I did see some improvement in my own sight reading ability, and also in my ability to hear intonation problems. I wish I could have stuck with it.

Subject H

I believe that this method does work, but requires certain things from the participant. First, I believe that my slow progress may be attributed to lack of quality time when practicing. Often I found that I was thinking about many other things, it was difficult to focus on the pitches. During the time when my schedule was more relaxed, I found improvement to be somewhat substantial. The time spent day to day does inevitably force one to create a tonal “memory bank” from which to draw on. My only regret is that I let this project take a back seat to so many other things. For this, I personally apologize and hope that my minimal participation will have aided in some small way—perhaps these comments are consistent with that of other participants who finished the study. Meeting with a partner, while an ideal situation, was infeasible for me at times. During the course of the study I had three different partners, all of whom I had difficulty meeting with. The individual practice sessions are not as effective as partner sessions. Once again, I did find this process to increase my tonal awareness for the pitches I studied. It is a mundane, yet effective, way of training the ear to really listen to the pitches.

Subject I

I believe that this can work. However, I wasn’t able to finish the program because of time constraints. I find it an interesting field and hope maybe to finish the program at a later date. In what I’ve done, I haven’t found it to help as far as using it in context.
Subject J
This did help develop my relative pitch sense. I was much more aware of intonation problems in Symphonic Choir. We did not practice weekends, nor did we practice much the week before our recital. I really do believe absolute pitch is acquired, especially after practicing. I feel that I am (maybe) halfway there. I was thinking of the implications for a choir of seniors, Spring Quarter, who have the taken the Burge course for two years. Imagine the perfect intonation and not having to give starting pitches. Thanks!

Subject K
Had class and partner scheduling conflicts, but I do think that the system works.

Subject L
I really didn’t have the proper kind of time to get the full benefit. I still feel it’s an advance course in relative ear training. If you didn’t work everyday on it, I feel you would regress. (This comes from my evidence of lack of time and sporadic work.) I got to Lesson Number Four.

Subject M
My main reason for dropping out of the study was time conflicts. Both regular course work and my recital demanded more time than I thought it would.

Subject N
I believe this study was beneficial to me in several ways. Although I haven’t acquired absolute pitch, my ability to recognize the pitches A and F# are about 90%. I have also learned to listen to the qualities of each individual pitch as opposed to always relating a pitch to another pitch or to a particular scale.

Subject O
It sharpened my ability to listen, hear, inner tones and express the overall music on the piano. I derive more pleasure from practicing as I am aware of more character of each individual tone. I also relax more and slow my playing down because I hate to have all those sounds (which I am becoming attached to) pass by so fast.

Subject P
My partner and I unfortunately were unable to finish all of the exercises in the designated time. The program was beneficial, however, I would have liked more “group” instruction. Example: meet three times a week for a week for a half-hour (like a class) and review and quiz a small group of us eight to ten at a time. That way the progress could have been more consistent and better monitored.
Subject Q

I had to drop absolute pitch training because of a heavy course load which caused a severe lack of time. I progressed to the seventh exercise and I did feel as if I made progress. It has improved my tonal memory greatly. I can now tell when someone is playing the same piece in a different key than from when I heard, maybe, several days before. I can tell when the tape played is playing a little slower or faster than "right."

Subject R
No comments.

Subject S
In the long run I'm sure the experiment has helped my ear. I usually can pick an A out when I hear it. Intonation problems seem to be much more obvious to me. When playing with groups it is easier to detect what I need to do to play in tune than it was before I did this experiment.

Subject T
I made it through. What's really strange is listening to the radio and naming the pitch. The study, by itself, was not quite enough to get it going, though. The relative pitch skills (especially relating notes to songs: Beethoven's 5th to remember G, etc.), were very helpful. However, I think the relative pitch, by itself, is not enough either. Together, the study and my previous training with relative pitch, I now have close to perfect pitch. This has been especially helpful in sight-reading skills (with or without) my instrument.

Subject U
The pace of our lessons on the average was about a lesson per three and one-half weeks. I got through the first four to five lessons before Christmas. I was up to ten or eleven by Spring, and did the rest after Spring Break. I've been working on fifteen for three weeks now. I think that it helped my relative pitch extensively. I also think that I've got a beginning sense of absolute pitch. I do much better when I hum; maybe because of vocal tension, maybe because I hear the pitch better. I think if I kept at it I could probably perfect it in time.

Subject V
I thought it was a good experiment, and designed well. As for my personal success, I believe that after reaching phase six much frustration was found on both my partners and my part. Without the motivation of success, our participation dropped off which only made it more difficult to advance to the next phase of the experiment. I really believe our own self-doubt prohibited us from succeeding.
Subject W
This experiment has been very helpful to me as a musician. I don’t think that I actually have perfect pitch, but I do feel relatively close. I would have liked for the group to get together more often to check on our progress. I personally should have been pushed more to keep on the ball. I have benefited from this study in my participation in Symphonic Choir and in accompanying different soloists. I feel that I have a better ear now and this will continue to help me in the future.

Subject X
As a trumpet player I find that transposition is a problem. I did feel that developing this took the same kind of practice that it takes on your instrument. This study will probably be useful in the future.

Subject Y
Sorry I didn’t complete my training because of conflicting time commitments.

Subject Z
My primary reason for discontinuing in the study was the large amount of time it took and the trouble my partner and I were having getting together. I also found that my partner was moving at a faster rate than I which was a little frustrating.