INFORMATION TO USERS

This material was produced from a microfilm copy of the original document. While the most advanced technological means to photograph and reproduce this document have been used, the quality is heavily dependent upon the quality of the original submitted.

The following explanation of techniques is provided to help you understand markings or patterns which may appear on this reproduction.

1. The sign or "target" for pages apparently lacking from the document photographed is "Missing Page(s)". If it was possible to obtain the missing page(s) or section, they are spliced into the film along with adjacent pages. This may have necessitated cutting thru an image and duplicating adjacent pages to insure you complete continuity.

2. When an image on the film is obliterated with a large round black mark, it is an indication that the photographer suspected that the copy may have moved during exposure and thus cause a blurred image. You will find a good image of the page in the adjacent frame.

3. When a map, drawing or chart, etc., was part of the material being photographed the photographer followed a definite method in "sectioning" the material. It is customary to begin photoing at the upper left hand corner of a large sheet and to continue photoing from left to right in equal sections with a small overlap. If necessary, sectioning is continued again — beginning below the first row and continuing on until complete.

4. The majority of users indicate that the textual content is of greatest value, however, a somewhat higher quality reproduction could be made from "photographs" if essential to the understanding of the dissertation. Silver prints of "photographs" may be ordered at additional charge by writing the Order Department, giving the catalog number, title, author and specific pages you wish reproduced.

5. PLEASE NOTE: Some pages may have indistinct print. Filmed as received.

Xerox University Microfilms
300 North Zeeb Road
Ann Arbor, Michigan 48106

THE DEVELOPMENT AND IMPLEMENTATION OF A RHYTHMIC ABILITY TEST DESIGNED FOR FOUR-YEAR-OLD PRESCHOOL CHILDREN

by

Robert H. McDowell

A Dissertation Submitted to
the Faculty of the Graduate School at
The University of North Carolina at Greensboro
in Partial Fulfillment
of the Requirements for the Degree
Doctor of Education

Greensboro
1974

Approved by

Walter L. Wehrer
Dissertation Adviser

APPROVAL PAGE

This dissertation has been approved by the following committee of the Faculty of the Graduate School at The University of North Carolina at Greensboro.

Dissertation Adviser Walter L. Wehmer

Committee Members

April 1, 1974
Date of Acceptance by Committee

MCDOWELL, ROBERT HARVEY. The Development and Implementation of a Rhythmic Ability Test Designed for Four-Year-Old Preschool Children. (1974) Directed by: Dr. Walter L. Wehner. Pp. 127

The purpose of this study was to develop a test designed to measure the rhythmic ability of four-year-old preschool children and to study the effects of training on posttest scores of the age-group. The forty-item test has four ten-item subtests which measure the following: the ability to differentiate whether or not two tempi are the same, the ability to accurately produce a given tempo, the ability to accurately reproduce a given rhythm pattern, and the ability to determine if two rhythms are alike.

The preliminary form of the test was revised to improve the item discrimination and item difficulty. The new test form was given to a total of 46 four-year-olds. The split-half test reliabilities ranged from .86 to .89 for these subjects. The Pearson product-moment formula in conjunction with the Spearman-Brown prophecy formula was used to compute the correlation. The test-retest reliability for the scores of 36 of these subjects who took the test twice was .45; however, there was an intervening variable for 24 of the subjects and the correlation yielded through the Pearson product-moment formula was not significant at .05.

Fourth-grade subjects (32) were given the old form of the test and also the Musical Aptitude Profile (MAP) of Edwin Gordon. This was done to validate the rhythm test with correlations between test scores on it and the MAP. The composite scores of the researcher's test correlated with the MAP scores as follows: MAP composite scores on "Rhythmic Imagery" (.65), "Rhythmic Imagery II"--meter (.61), "Rhythmic Imagery I"--tempo (.60), and the composite for the whole MAP (.51).

These correlations, derived from the Pearson product-moment formula, were all significant at or beyond .001. Validity was also studied by correlating the scores of 10 kindergarten subjects with their music and classroom teachers' rankings of their "musical rhythmic ability." The paired comparison technique was used for the rankings and the Spearman rank-order formula for the correlations. The correlations were .78 (music teacher) and .66 (classroom teacher).

The revised test was given as a pre- and posttest to 36 four-year-olds who were divided into three groups. One group had no contact with the researcher between the pre- and posttest. Another group received 20 sessions of contact with the researcher in the month between the tests. The sessions involved nonmusical activities. A third group received 20 sessions of rhythmic training. The researcher hypothesized that the group receiving the training would show a significant increase in their posttest scores when compared to the performances by the other two groups. An analysis of covariance yielded an F ratio which was not sufficient to reject the null hypothesis at .05. Hypothesis 2 stated that subjects having contact, but not the training, would improve more on the posttest than the subjects having no contact. The null hypothesis was accepted since the group not having contact had more improvement. This was inconsistent with this hypothesis. The third hypothesis, that sex would not significantly affect performance on the test, was tested using analysis of variance. The analysis yielded an F ratio which was not significant at .05, and the hypothesis was accepted.

It was concluded that one month of rhythmic training did not significantly affect the rhythmic ability of four-year-olds.

ACKNOWLEDGEMENTS

The writer expresses his appreciation to Dr. Walter L. Wehner for his willing and valuable guidance during the course of this investigation and for his personal interest during the writer's graduate study at The University of North Carolina at Greensboro.

Thanks are also given to Dr. Harold F. Abeles for advice concerning some of the statistical aspects of this study.

The writer expresses appreciation to the students, teachers, administrators, and others who helped in the research from Hester's Creative Schools, The Fairfax County Public Schools, The Montgomery County Child Day Care Association, and The Prince George's County Public Schools. Without their splendid cooperation this project could not have been possible.

The writer also acknowledges his gratitude to Gail Edgerton McDowell for her assistance and encouragement throughout this investigation.

TABLE OF CONTENTS

APPENDIX

LIST OF TABLES

TABLE | Page

CHAPTER I

INTRODUCTION AND PROBLEM

In terms of presently available data, one of the most neglected areas in music research is early childhood education. Early childhood education, also referred to as nursery school or preschool, has been defined by Sigel, who states:

> A preschool environment as considered in this paper refers to a group setting away from home for children under public school age (i.e., children ranging in age from two to five) under the aegis of teachers.[1]

Many authorities concerned with the learning process feel that what happens to a child during the period of early childhood can strongly affect that child in years to come. Bloom states:

> . . . intelligence, the capacity to learn, grows as much during the first four years as in the next thirteen. After age seventeen intelligence continues to develop but at a comparatively slow pace. Failure to develop proper learning patterns in the pre-primary years is likely to lead to continued failure or near failure throughout the remainder of the individual's school career.[2]

Governmental projects such as Project Head Start indicate a national concern that children need appropriate experiences early in life which will enhance their mental growth, cultural awareness, and

---

[1]Irving E. Sigel, "Developmental Theory and Preschool Education: Issues, Problems and Implications," Early Childhood Education, Seventy-fifth Yearbook of the National Society for the Study of Education, Part II (Chicago: The University of Chicago Press, 1972), p. 15.

[2]Benjamin S. Bloom, Allison Davis, and Robert Hess, Compensatory Education for Cultural Deprivation (New York: Holt, Rinehart and Winston, Inc., 1960), p. 16.

human potential. Akers presents a strong argument for the importance of the formative years to human growth and development. He writes:

> The recent concentration of investigation and study by psychologists, educators, pediatricians, psychiatrists, anthropologists, nutritionists, and others point clearly in one direction. The child's earliest years are the time of most rapid physical and mental growth. At no other period in his life is he so susceptible and responsive to positive environmental influences which enhance and expand his development. Environmental influences, if of a sterile or destructive nature, may have negative effects on his intelligence, his motivation and ability to learn, his concept of himself, his relationship with others, and on his later health.[3]

This writer makes the basic assumption that early childhood is a very important stage of growth and development, and he believes that this area needs more research and investigation.

The relationship between heredity and environment and their effects on human growth and development have been points for strong disagreement among many people of the scientific community. Many believe that heredity plays a more important role; others believe that environment is the stronger contributor to human development. The controversy has been particularly intense in the area of music aptitude. A researcher who has been linked with the idea that music aptitude is more or less innate is Carl Seashore. Seashore writes:

> On the basis of our experiments in measuring these sensory capacities, we find that the basic capacities, the sense of pitch, the sense of time, the sense of loudness, and the sense of timbre are elemental, by which we mean that they are largely inborn and function from early childhood.[4]

---

[3]Milton E. Akers, "Prologue: The Why of Early Childhood Education," Early Childhood Education, Seventy-fifth Yearbook of the National Society for the Study of Education, Part II (Chicago: The University of Chicago Press, 1972), pp. 2-3.

[4]Carl E. Seashore, Psychology of Music (New York: McGraw-Hill Book Company, Inc., 1938), p. 3.

Robert Lundin is among those authorities who believe that musical responses are linked closely with environment. He writes:

> Musical responses, then, are learned. The learning may occur through a casual contact with various musical stimuli, or it may occur quite deliberately through training by qualified instructors. No one is born gifted with any "powers" which will destine him to be a genius, musically speaking. This is not to deny that some persons are more biologically predisposed to respond to musical stimuli than others.[5]

Probably neither Lundin nor Seashore represents the extreme position in its pure form. From the previous quotes of Seashore and Lundin respectively, the phrases "largely inborn" and "more biologically predisposed" appear to soften their positions somewhat.

Farnsworth aptly describes the controversy. He writes:

> The present-day formulation of the nature-nurture relationship is not one which would have appealed to the extremists of the 1920's and 1930's. Whether they were hereditarians or environmentalists, the older theorists blinded themselves to the obvious in their attempts to maintain their one-sided positions. It is now clear that neither nature nor nurture alone can make a musician. Both must be present before musical and other abilities can emerge.[6]

Glenn and Turrentine also point out that a child's responsiveness to music depends on both environment and heredity. They write, "Suffice it to say, if maximum conditions of nature and nurture are not present, maximum growth and development will not take place."[7]

---

[5]Robert W. Lundin, An Objective Psychology of Music (New York: The Ronald Press Company, 1967), pp. 8-9.

[6]Paul R. Farnsworth, The Social Psychology of Music (Ames, Iowa: The Iowa State University Press, 1969), p. 156.

[7]Neal E. Glenn and Edgar M. Turrentine, Introduction to Advanced Study in Music Education (Dubuque, Iowa: Wm. C. Brown Publishers, 1968), p. 48.

Some researchers believe that music aptitude is established at an early age; however, they often disagree as to whether this is a product of environment or heredity. Seashore states:

> The apparently complex forms of sensory capacities also tend to be elemental to a considerable degree; that is, the young child has the sense of tone quality, of volume, of rhythm, and the sense of consonance long before he begins to sing or know anything about music.[8]

More specifically, Seashore believes that music aptitude is permanently fixed by age ten. He says, "We can measure these capacities reliably by the age of ten in the normal child; and this measure is likely to stand, except for the numerous vicissitudes of life which may cause deterioration."[9]

Gordon presents evidence to support Seashore's claim. Based on studies of the Musical Aptitude Profile[10] (MAP) by Fosha[11], Gordon[12], and Tarrell[13], Gordon states, " . . . scores of both fourth-grade students and older students remain stable even after they have been exposed to musical practice and training."[14]

---

[8]Seashore, Psychology of Music, p. 3. [9]Ibid.

[10]Edwin Gordon, Musical Aptitude Profile Manual (Boston: Houghton Mifflin Company, 1965).

[11]Leon Fosha, "A Study of the Validity of the Musical Aptitude Profile" (unpublished Ph.D. dissertation, University of Iowa, 1960).

[12]Edwin Gordon, A Three-Year Longitudinal Predictive Study of the Musical Aptitude Profile, Vol. V: Studies in the Psychology of Music (Iowa City, Iowa: University of Iowa, 1968).

[13]Vernon Tarrell, "An Investigation of the Validity of the Musical Aptitude Profile," Journal of Research in Music Education, XIII (Winter, 1965), 195-206.

[14]Edwin Gordon, The Psychology of Music Teaching (Englewood Cliffs, N. J.: Prentice-Hall, Inc., 1971), p. 5.

Gordon attributes music aptitude to both hereditary and environmental influences. He sums up his position on music aptitude as follows:

> Musical aptitude is a product of innate potential and early environmental influences. It is normally distributed among students of all ages. The main dimensions of musical aptitude are rhythmic, tonal, and aesthetic-interpretive. Although musical aptitude fluctuates throughout the primary grades, it becomes impervious to practice and training at about age ten.[15]

For purposes of this study, the author takes the position that a child's musical aptitude depends on both heredity and environment. Further, there is evidence which indicates that a child's environment seems to be crucial in the development of his musical aptitude; this is especially true for the period of conception to about age ten.

There are many unsolved problems relating to the development of music aptitude in young children such as: Is it a gradual process? Does it occur in spurts? What experiences help the child reach a maximum music aptitude? What experiences thwart the growth of musical aptitude? In order for researchers to begin answering these questions some basic research tools are needed. The music aptitude tests discussed by Lehman[16], Whybrew[17], and others are not designed for use with preschool children. The Measures of Musical Ability[18] by Bentley is

---

[15]Ibid., p. 7.

[16]Paul R. Lehman, Tests and Measurements in Music (Englewood Cliffs, N. J.: Prentice-Hall, Inc., 1968).

[17]William E. Whybrew, Measurement and Evaluation in Music (2nd ed.; Dubuque, Iowa: Wm. C. Brown Company Publishers, 1971).

[18]Arnold Bentley, Measures of Musical Ability (New York: October House, Inc., 1966).

the only battery with adequate standardization which can be used to measure the musical aptitude of children as young as seven. Since the tests of Bentley, Gaston, Gordon, and others are normally taken via the use of answer sheets, they are inappropriate for use with preschool children who, for the most part, can neither read nor write with enough facility to take these tests.

When this study began, if an experimenter wished to measure the rhythmic ability of four-year-olds he had to modify an existing test or create a new test. The writer carefully studied the musical aptitude tests which are adequately standardized and in print. The researcher also became acquainted with some of the behavior patterns of three- and four-year-old children through the following: observations in a preschool setting, informal interviews with preschool children, interviews with nursery school teachers, interviews with nursery school administrators, and perusal of literature about preschool children. After this research, the author concluded that he could not modify any of the currently available music aptitude tests for use with preschool children without destroying the original format, logic, and internal structure of the test.

In this study the author sought to develop data on a pilot test battery which he had constructed to measure the rhythmic ability of four-year-old children. Also, using the pilot battery he sought to determine what effects training would have on posttest scores. The writer hoped that this research may lead to a standardized test or tests which can be used as tools for more research.

Children of the age of four were selected as the most appropriate for this study. In a preliminary investigation, it was discovered that three-year-olds, as a group, could not give responses to preliminary test items which could be recorded by the researcher with much consistency or reliability. This is not to say that children age three do not have measurable responses to stimuli. The writer simply states that he was unable to construct a test which he felt was suitable for use with three-year-old or younger children.

The rhythm test to be used in this study is designated as a test of the rhythmic "ability" of four-year-olds. Since it seems quite possible that the music aptitude of a child is in a state of flux until about the age of ten, then, to use the term "aptitude" would not seem appropriate. In a discussion of "talent," "capacity," "ability," and "aptitude," Farnsworth writes:

> The term "ability," suggesting the power to act but indicating nothing about the heritability or congenitalness of inferred potentiality, is the broadest and safest of all these terms. As we shall soon see, nature and nurture invariably function jointly, and it is erroneous to say that any act is the sole result of either the one or the other.[19]

Seashore and many others have recognized the difficulties involved in testing young children. In a discussion of the measurement of musical capacities in children, Seashore states, " . . . the exhibition of these capacities is limited by the child's ability to understand or apply himself to the task."[20]

---

[19]Farnsworth, The Social Psychology of Music, pp. 151-152.

[20]Seashore, Psychology of Music, p. 3.

One of the values of this study would be the addition of more information to the literature about four-year-olds and their responses to rhythmic stimuli in a controlled situation. This would be accomplished through a presentation and analysis of data collected from the subjects' test scores and their responses to the various items on the researcher's test.

Another value of the study would be the information gained concerning testing procedures which can be successfully used with this age-group in investigating responses to musical stimuli. Also, an investigation of the effects of training on posttest scores might give an indication of the stability of rhythmic ability at this age.

All good tests have satisfactory levels of reliability and validity. The writer collected data in preliminary research with four-year-olds which gave an indication that a preliminary form of the rhythm test had a fairly high test reliability. The researcher proposed to use fourth-grade subjects to determine the validity of his test by discovering how scores on his test correlate with scores of the same group of subjects on the MAP. Also, the researcher proposed to see how well the rhythm test scores correlated with teacher rankings of rhythmic ability.

Most authorities agree that music aptitude or ability has more than one facet or component. Many believe that music aptitude should include the measurement of rhythmic response in some form or other. The results of this study may be helpful to researchers seeking to construct a music test for preschool children which might include measures

of rhythmic, melodic, harmonic, aesthetic, or other factors.

The writer made the following hypotheses about the pilot test battery which were tested:

1. The test scores of four-year-old experimental subjects will significantly improve after training.

2. The scores of the control subjects having contact with the researcher between testings will show a larger mean increase on the posttest than the subjects not experiencing the contact; however, the larger increase will not be a significant one.

3. Among the subjects in the experiment, there will be no significant difference in the performance on the test which can be attributed to sex.

The above were stated as null hypotheses in order to statistically test the results as follows:

1. There will be no significant differences found between pre- and posttest scores of subjects receiving training and those not.

2. The scores of the control subjects having contact with the researcher between tests and the subjects not experiencing contact will show no significant differences in test scores.

3. Among the subjects in the experiment, there will be no significant difference in the performance on the test which can be attributed to sex.

CHAPTER II

REVIEW OF THE LITERATURE

Introduction

The history of the preschool movement in the United States dates from around 1920 when the first formal preschool centers were established. Concerning the early development of the preschool, Evans writes:

> With the gathering momentum of the child guidance movement in the late 1920s, nursery schools began to flourish. Several important centers for child study operated model programs, including the Gessell Child Guidance Nursery at Yale University, the Merrill-Palmer Institute in Detroit, Teachers College (Columbia University), and the Iowa Child Welfare Research Station at the University of Iowa. Emerging legislation during the great depression, subsumed under Franklin Roosevelt's WPA program, created a federal nursery school sponsorship.[1]

In recent years preschool education has grown in both the number of children involved and the number of preschools. Concerning this growth, LaCrosse writes:

> The United States Census Bureau reports that in 1965 one of ten children three to four years old were in some type of formal preschool program. In 1970 that figure was one in five. Movement in the field of early education in the past five years has been like an avalanche. There is a large gap between our need to know and the available information.[2]

---

[1]Ellis D. Evans, Contemporary Influences in Early Childhood Education (New York: Holt, Rinehart and Winston, Inc., 1971), pp. 12-13.

[2]Robert E. LaCrosse, ed., Early Childhood Education Directory (New York: R. R. Bowker Company, 1971), p. vii.

Reviews of early childhood research having information on music have appeared at rather sporadic intervals since 1932. In that year Williams[3] reviewed the research literature dating from the 1920's to the early 1930's. A similar review of Jersild[4] (1939) overlaps and extends the review of Williams. More recently, Fowler discusses the literature relating to musical ability in young children in an article on infants and young children.[5]

The author will present below a review of the literature pertaining to research done with preschool children in the area of musical response. These studies, reports, and articles often include reports of research done with kindergarten and elementary subjects as well as preschool subjects. Some of the studies involve several factors which often involve and overlap each other. For clarity, the writer has grouped these studies under seven headings as follows: research into the singing response, research into the use of musical instruments, research using operant training, research into the rhythmic response, development of ranking and rating scales, research into home environment and musical response, and modification of existing tests for research.

---

[3]Harold M. Williams, "Studies in the Measurement of Musical Development," in The Measurement of Musical Development, Vol. VII, No. 1: University of Iowa Studies: Studies in Child Welfare (Iowa City, Iowa: University of Iowa, 1932), pp. 9-31.

[4]Arthur T. Jersild, "Music," Child Development and the Curriculum, Thirty-eighth Yearbook of the National Society for the Study of Education, Part I (Bloomington, Illinois: Public School Publishing Company, 1939), pp. 135-151.

[5]William Fowler, "Cognitive Learning in Infancy and Early Childhood," Psychological Bulletin, LIX (March, 1962), 135-137.

Research into the Singing Response

Jersild and Bienstock were pioneers in the investigation of the singing of preschool children. The subjects included in one study were 48 three-year-old children. An experimental group of 18 subjects received training in interval and pitch reproduction over a six-month period. Each child in the experimental group was observed for 100 minutes while in a free-play situation, and observers kept records of the spontaneous vocalizations of the subjects. The researchers concluded that training markedly improved the singing ability of the experimental group; however, the researchers were cautious not to imply that any change in native ability had taken place. The researchers also discovered that children's voices were placed somewhat lower than was commonly believed.[6]

A similar study was made by the same researchers about three years later. No control group was used in this study of 23 subjects who ranged in age from three to five and one-half years. The subjects' singing responses were measured to see how well they could sing intervals and pitches. The findings indicate that children can sing a wide range of pitches and intervals at an early age. Also, Jersild and Bienstock found that chromatic intervals were no harder for the children to sing than unaltered diatonic intervals.[7]

---

[6]Arthur T. Jersild and Sylvia F. Bienstock, "The Influence of Training on the Vocal Ability of Three-Year-Old Children," Child Development, II (December, 1931), 289-291.

[7]Arthur T. Jersild and Sylvia F. Bienstock, "A Study of the Development of Children's Ability to Sing," Journal of Educational Psychology, XXV (October, 1934), 481-497.

The results from both studies of Jersild and Bienstock indicate that a child's singing ability improves after training. This writer points out that the studies measured the ability to sing pitches and intervals. As such, the studies were not measures of the ability of children to sing songs.

Williams conducted a two-part experiment with subjects ranging from two and one-half to six and one-half years old. In the first part, a subject's success in reproducing tones and short musical phrases was rated by two observers. The agreement between the observers' ratings led Williams to believe that this was a "reliable and convenient" way of measuring the vocal abilities of these children.[8]

In the second part of the experiment, subjects were taught select songs during a regularly scheduled music class. The subjects were 41 four- and five-year-olds. A dictaphone was used to record the singing of each child after he had been exposed to the songs for a school year. The data indicated marked differences among four-year-olds in the ability to learn simple tunes. Williams attributed this to maturation. He also found that some children make gross errors in singing some intervals but return to the tonic at the end of a phrase.

Hattwick conducted a series of experiments in Iowa using 3,902 subjects. The subjects ranged in age from three years and six months to twelve years and four months. In one experiment with 126 preschool and elementary school subjects he measured the "voco-motor" ability

---

[8]Williams, "Studies in the Measurement of Musical Development," p. 79.

using accuracy of interval singing. Stimuli were produced by metal bars having resonator tubes. Hattwick found it a difficult task to get some of the older subjects to respond. Over fifty percent of the fifth-grade subjects were not testable even after training.[9] The researcher presents a manual of instructions with norms to be used with his test. He does not recommend the test for children who are younger than second-grade because of the lower reliability data of these test scores. Subjects older than fourth-graders cannot be tested using the instrument.

Hissem made an eight-month study of 27 subjects ranging in age from 21 to 54 months when the experiment began. The experimenter worked with each child taking into account individual differences. Generally, the training periods involved the experimenter and subject working with a dinner chime based on the tonic chord of C Major. The experimenter concluded that children respond well to this type of training, and there was systematic improvement in pitch matching, tonal discrimination, and rhythmic discrimination.[10]

In another study , Hattwick researched the pitch levels and singing ranges used by preschool, first-grade, and second-grade subjects. The 95 subjects were divided into two different groups. One group sang 47 different tunes at any pitch level they selected. The other group

---

[9]Melvin S. Hattwick, "A Genetic Study of Differential Pitch Sensitivity," in The Measurement of Musical Development II, Vol XI, No. 2; University of Iowa Studies: Studies in Child Welfare (Iowa City, Iowa: University of Iowa, 1935), p. 21.

[10]Irene Hissem, "A New Approach to Music for Young Children," Child Development, IV (December, 1933), 309.

(37 subjects) was given 48 practices on one song starting on the same pitch each time. Hattwick found that the group which had been exposed to the one song had a significantly lower pitch-level when subjects sang the song individually at the pitch each subject selected. Also, the mean pitch-range used by the children when they sang alone was significantly lower than the pitches of songs found in current song books appropriate for the age-levels.[11]

Updegraff, Heiliger, and Learned studied the effects of training on the musical interests and singing of subjects ranging from three to five years old. The subjects numbered 16 three-year-olds, 14 four-year-olds, and 36 five-year-olds.[12] The subjects at each age-level were divided into two groups according to preliminary tests of singing, tests of rhythm, interest in music, and musical background.

The researchers discovered that the training program enabled the experimental group to improve their reproduction of pitches, intervals, and musical phrases. The subjects in the control group improved only slightly or maintained the same level of ability on a posttest. This data seems to substantiate the earlier findings of the experimenters Jersild and Bienstock.

---

[11]Melvin S. Hattwick, "The Role of Pitch and Pitch Range in the Singing of Preschool, First Grade, and Second Grade Children." Child Development, IV (December, 1933), 290.

[12]Ruth Updegraff, Louise Heiliger, and Janet Learned, "The Effect of Training Upon the Singing Ability and Musical Interest of Three-, Four-, and Five-year-old Children," in Studies in Preschool Education I, Vol. XIV, No. 1: University of Iowa Studies: Studies in Child Welfare (Iowa City, Iowa: University of Iowa, 1938), p. 92.

Drexler also studied the singing response. The subjects in her study had a median age of four years and five months. Drexler taught the 23 subjects a melody. Using the Ediphone (a forerunner of the Dictaphone) she recorded the singing of each child after the training period. The tones sung by the children were later transcribed into musical notation and plotted against the correct notation on a graph. Drexler found that the ability to "carry a tune" (sing the correct pitches) increased with age, and she found that the differences in this ability were most significant between the ages of three to four and five to six. Drexler also found that smaller descending intervals were easier for the children to accurately sing than ascending ones. There were no observable differences which could be attributed to sex.[13]

Smith reports on the effectiveness of preschool music classes which were taught by a music specialist. The subjects whose data were used in this study number 13 three-year-old and 16 four-year-old children.[14] This study differs from most of the previous studies in that the subjects were trained in fairly large groups. The music activities of the children were varied; however, singing ability is the only one of the activities reported on in this study. The subjects were exposed to an orientation period prior to the experiment. In this period various musical experiences were conducted by the music specialist. After

---

[13]Edith N. Drexler, "A Study of the Development of the Ability to Carry a Melody at the Pre-School Level," Child Development, IX (September, 1938), 331.

[14]Robert Barton Smith, "A Study of the Effect of Large-Group Vocal Training on the Singing Ability of Nursery School Children" (unpublished Ed.D. dissertation, University of Illinois, 1960), p. 89.

the orientation, the subjects in both groups were given a vocal test. On the basis of data from the test and findings of previous researchers, Smith designed a two-semester curriculum intended to help the subjects improve their singing ability. After exposure to the curriculum, the three-year-olds could sing the range of a sixth accurately. This group did not respond to training designed to help them sing accurately in a wider range. The four-year-olds accomplished the same results in one semester, and they improved their range in the second.

Research into the Use of Musical Instruments

Colby experimented with the ability of subjects to perform certain musical tasks on a tin fife. The subjects were 16 children (eight of either sex) who were between three and one-half and four and one-half years old.[15] The purpose of the experiment was an application of preliminary training to the learning and performance of 25 melodic patterns by the subjects. Colby found that subjects with short fingers had difficulty. Also, he found it time-consuming and difficult to teach the melodic patterns to the subjects. Colby concluded that instrumental training in preschool did not yield as good results as vocal training could. To this writer, Colby's procedures seemed, perhaps, too formal and too highly structured for this age-group.

---

[15]Martha G. Colby, "Instrumental Reproduction of Melody by Preschool Children," Journal of Genetic Psychology, XLVII (December, 1935), 414.

Research Using Operant Training

Fullard used operant training to improve the aural discriminative powers of 10 preschool children. Through positive reinforcement with M&M candy, the subjects were trained to identify by sight and sound the violin, clarinet, violoncello, flute, viola, and French horn.[16] Through a comparison of pretest and posttest scores, Fullard concluded that learning had taken place, and operant training of this type could be used beneficially with preschool subjects.

Research into the Rhythmic Response

One of the earliest and best controlled experiments on rhythmic ability was conducted by Heinlein. Eight subjects, ranging in age from three to five, marched to a composition played on a player piano which was electrically operated. The subjects were standing on an electrified platform. Each subject had a foot stirrup attached to his right shoe. When the stirrup touched the platform it closed an electrical circuit. A series of kymograph pens were used in the experiment to objectively record the marching responses. A group of adult subjects pushed telegraph keys which recorded what the adults thought was the exact instant that a subject's foot touched the platform. Heinlein concluded that subjective observations, as represented by the adult subjects' responses, were unreliable. None of the adults were music teachers or professional musicians. Only two of the children were able to

---

[16]William G. Fullard, Jr., "Operant Training of Aural Musical Discriminations with Preschool Children," Journal of Research in Music Education, XV (Fall, 1967), 204.

synchronize their marching with the music to a degree which Heinlein considered to be very successful.[17]

Williams experimented with a subject's ability to tap synchronously with mechanically-produced periodic patterns of sound. A modified version of the Seashore motor apparatus was used. Children responded to a stimulus by tapping a small hammer on a small piece of metal. The tapping closed a circuit which registered a mark on a revolving disc. The experiment involved two test series. One series was a presentation of stimuli of constant periodicity. The other series involved periodic stimuli, but there was a variation in the rate of the stimuli from one test item to the next. The subjects were divided into two groups. There were 131 children in the group who ranged from four to eight years in age, and there were 82 in the other group who were from five to twelve years old. The results of this study indicate that children are able to respond more accurately to the modified Seashore apparatus as their age increases; however, Williams found that there was more overlap in ability as children became older. Subjects were more successful in their responses when the stimuli were between .40 and .50 times per second apart.[18]

Williams gave two motor tests to 30 selected subjects. These subjects ranged in age from four and one-half to six and one-half years

---

[17]Christian Paul Heinlein, "A New Method of Studying the Rhythmic Responses of Children Together with an Evaluation of the Method of Simple Observation," Journal of Genetic Psychology, XXXVI (June, 1929), 205-222.

[18]Williams, "Studies in the Measurement of Musical Development," pp. 32-66.

old. The tests were developed at the University of Iowa by Hicks[19] and Wellman.[20] The reliability of Hicks' test ranged from .77 to .86 for the age group presently under discussion. The Wellman test had a reliability of .85 to .90 for a similar age-group. The correlation of mean error scores of Williams' test with Hicks' test was about .36. Correlation with the Wellman test was around .25.[21]

Jersild and Bienstock also experimented with the ability of children to keep time to a musical stimulus. After a preliminary experiment with preschool subjects, the researchers concluded that subjective measures were not accurate enough for their purposes. The researchers developed more objective measures of clapping and walking responses using motion pictures. These pictures simultaneously recorded the responses of the subjects together with the strong beats of the music via light flashes. The light source and a one-second clock were located nearby the subject so all three would be filmed together. The 95 subjects used in the study ranged in age from two to five; however, all subjects did not participate in all of the divisions of the study. The reliability data based on two administrations of the test ranged from about .56 to .90 for the various age-levels. The scores tended to increase as the subject's age increased. No notable differences were

---

[19]James A. Hicks, The Acquisition of Motor Skill in Young Children, Vol. IV, No. 5: University of Iowa Studies: Studies in Child Welfare (Iowa City, Iowa: University of Iowa, 1931).

[20]Beth Wellman, The Development of Motor Co-ordination in Young Children, Vol. III, No. 4: University of Iowa Studies: Studies in Child Welfare (Iowa City, Iowa: University of Iowa, 1926).

[21]Williams, "Studies in the Measurement of Musical Development," p. 55.

discovered by the experimenters in reference to sex and test scores at a given age-level. Correlations of intelligence scores and rhythmic performance was variable--most were positive. The data yielded a correlation coefficient of about .30 when singing ability (number of tones a child could accurately reproduce) was compared with the scores on the tests. Correlations between the "clapping" and "walking" scores were about .80.[22]

The researchers concluded that rhythmic ability can be attributed largely to maturity. This study is well-constructed and well-controlled, in the opinion of this writer; however, one may conclude that this was more of a study of the subjects' abilities to respond to a pulse rather than a study of rhythmic ability. Responding to a pulse is probably only one facet of rhythmic ability.

In 1936, Wight performed an experiment on the motor and "rhythmization" abilities of handicapped children of the County Home for Convalescent Crippled Children in Chicago. The subjects included a few preschool children, and the children's ages ranged from 57 to 187 months. The ability to "rhythmize" was measured by the subjects' ability to reproduce rhythmic patterns produced by strokes of a magnetic hammer wired to contacts on a revolving disc. Motor skill was measured by tapping tests. The subjects alternately tapped two keys in one of the tests; in the other test they consecutively tapped six keys. Speed in tapping was the criterion for measuring motor skills. Wight's find-

---

[22]Arthur T. Jersild and Sylvia F. Bienstock, Development of Rhythm in Young Children, Child Development Monographs, XXII (1935), 87-93.

ings yield little data pertaining directly to preschool children. Wight arrived at the following conclusions:

1. There are individual differences in ability to rhythmize.
2. Rhythmization and intelligence are both related to motor coordination.
3. Rhythmization is subject to improvement through both specific and general training, no matter what the initial level of ability may be.[23]

Christianson made a two-year study of the rhythmic movements of young children. The 47 subjects ranged in age from two years to six years and six months. Christianson and three assistants spent a year developing a rating scale and methods of recording the bodily and verbal responses of children to a musical stimulus. The rating scale measured the degree of responsiveness in five categories: synchronization with the rhythm of a piece of music; social and emotional responses in the presence of a musical stimulus; dancing behavior; use of rhythmic movements to enhance dramatic expression; and requests or comments by the children pertaining to musical activities. For purposes of the study, 85 musical selections were picked by a panel of experts in early childhood education, dance, and music education. All of the selections were markedly rhythmic. Each subject was observed for a four-week period. A child's responses were noted by two observers who were trained in music.[24] The experimenter considered her experiment to be

---

[23]Minnie G. Wight, "The Effect of Training on Rhythmic Ability and Other Problems Related to Rhythm," Child Development, VIII (June, 1937), 171.

[24]Helen Christianson, Bodily Rhythmic Movements of Young Children in Relation to Rhythm in Music, No. 736; Contributions to Educa-

only a tentative one, but she felt that her rating scale would be useful because of a lack of more objective means of measurement. She found that subjects made progress (made higher ratings) as their age-level increased. This study might also be included in the following group.

Development of Ranking and Rating Scales

At least two studies involved, primarily, the rating or ranking of children on musical behaviors. Vance and Grandprey sought to develop an objective method of ranking nursery school and kindergarten subjects since they felt that the Seashore test could not be used with these children because of their immaturity. The rankings were based on items which fall into the following categories: a child's response to musical stimuli in a kindergarten or preschool setting, the musical aspects of the child's home environment, a child's ability to beat in time with music, a child's ability to reproduce intervals vocally, and a child's ability to imitate rhythmic patterns by beating on a small triangle.[25] The researchers hoped to make a correlational study of their ranking scores and scores on the Seashore tests when the subjects were old enough to take the Seashore measures.

One of the more important findings of the study was a correlation of .62 between a child's home environment and his responsiveness to music as measured by the rankings. The correlation of rhythmic

---

tion (New York: Bureau of Publications, Teachers College, Columbia University, 1938), pp. 30-32.

[25]Thomas F. Vance and Medora B. Grandprey, "Objective Methods of Ranking Nursery School Children on Certain Aspects of Musical Capacity," Journal of Educational Psychology, XXII (November, 1931), 578.

capacity and age was .03 in the kindergarten group; however, nursery school scores correlated with age as high as .37. The researchers felt that the maturity of the nursery school children was being measured more than their musical capacity.[26]

About five years after the previous study, Seltzer reported the development of a rating scale to be used with preschool children. The scale was devised to measure the rhythmic and singing development of children, because, as the researcher believed, the Kwalwasser-Ruch and Seashore aptitude tests were not appropriate for preschool children.[27]

The researcher devised two lists of statements pertaining to the musical behaviors of preschool children. One list was to be used for rating subjects on singing ability, and the other was to be used for rating rhythmic ability. A panel of judges was used to weight the items. In final form, the rhythmic scale had 44 items for rating rhythmic ability and 42 items for rating singing ability. On the scale from one to eleven, one represented " . . . the greatest possible lack of progress in singing or rhythmic development, or, at eleven, the greatest possible development, with the midpoint at six."[28] The researcher believed that, when a sufficient number of children had been rated by these scales, then, a percentile table could be devised to interpret scores.

---

[26]Ibid., p. 583.

[27]Serkphine Seltzer, "A Measure of the Singing and Rhythmic Development of Preschool Children," Journal of Educational Psychology, XXVII (September, 1936), 417.

[28]Ibid., p. 419.

Research into Home Environment and Musical Response

Many studies have dealt with home environment and musical response as a by-product of an investigation. Five studies deal almost exclusively with this area. There are studies by Shull[29], Broadhead[30], Reynolds[31], Kirkpatrick[32], and Shelton.[33] There is much controversy among these researchers concerning the definition of musical and unmusical environments. Also, there is much deviation in the methods used by the various researchers to determine a child's musical ability. The study of Shull used classroom teachers to determine the ranking of subjects according to musicality. Broadhead developed her own test to measure musical ability. Reynolds and Kirkpatrick equated musical ability with the ability to sing. Shelton used elementary music teachers to rate the children.

None of the studies presents data which can unquestionably show a relationship in degree or type between a child's home environment and

---

[29]Dorothy Shull, "A Study of the Influence of the Musical Environment in the Home on the Musicality of Selected Kindergarten Children" (unpublished Master's thesis, Northwestern University, 1953).

[30]Mary Caryl Broadhead, "Musical Opportunities in the Home and Their Relationship to the Musical Achievement of Kindergarten Children" (unpublished Master's thesis, Cornell University, 1958).

[31]George E. Reynolds, "Environmental Sources of Musical Awakening in Pre-School Children" (unpublished Ed.D. dissertation, University of Illinois, 1960).

[32]William C. Kirkpatrick, Jr., "Relationships Between the Singing Ability of PreKindergarten Children and Their Home Musical Environment" (unpublished Ph.D. dissertation, University of Southern California, 1961).

[33]John Stanley Shelton, "The Influence of Home Musical Environment Upon Musical Response of First-Grade Children" (unpublished Ed.D. dissertation, George Peabody College for Teachers, 1965).

his musical ability or responsiveness to music.

Modification of Existing Tests for Research

Seashore's music aptitude test has been widely used with older children and adults. At least two experimenters have used a modified form of the test with preschool subjects. McGinnis modified the subtests for use with subjects ranging from 41 to 59 months. She did this because the original test was too long, children had difficulties in understanding some of the terms used in the test, the test items were uninteresting to young children, and some subtests were inappropriate for use with preschool children.[34]

The researcher selected only the tests of intensity, pitch, and consonance for administration. She divided each test into two parts to keep the subjects from becoming fatigued or losing concentration. She also substituted "loud" and "soft" for Seashore's "weak" and "strong;" "baby bear" and "daddy bear" for "high" and "low;" "pretty" and "ugly" for "better" and "worse."[35]

The experimenter tried to make a game out of taking the tests. For example, in the pitch test, the subject was to tell who spoke the last note, the baby (higher note) or the daddy (lower note). The three tests were administered twice to each subject individually. McGinnis compensated for omitted responses by taking the percent of the correct

---

[34]Esther McGinnis, "Seashore's Measures of Musical Ability Applied to Children of the Pre-School Age," American Journal of Psychology, XL (October, 1928), 620-621.

[35]Ibid.

responses from the total number of the responses of each child. McGinnis found that the reliability data from her modified form of the Seashore subtests compared favorably with data gathered by others who had done reliability studies on the Seashore test. There was a small but positive correlation between chronological age and the scores on the test as modified by McGinnis. There was also a small correlation between test scores and mental age. McGinnis concluded that performance on the test did not depend on chronological or mental age.

A later study by Friend was based on the same three tests which McGinnis used, and the Friend study appears to be closely modeled after that of McGinnis. Friend used 42 subjects ranging in age from 51 to 74 months.[36] As in the McGinnis test, terms such as "loud" and "soft" were substituted for "strong" and "weak." Each subject was given the subtests twice. In most cases the second administration followed the first by less than a week's time. The experimenter administrated the three tests at one sitting. One might question this procedure, especially with young children. Each subject was presented with 200 stimuli at each sitting.

The reliability coefficient of scores between the two administrations was .778. Friend had both parents of each child and the child's teachers rate him on a five-point scale of musical ability. These correlations were relatively low; the parent ratings correlated higher (.264) with the test scores than did the teachers' ratings. The parents

---

[36]Ruby S. Friend, "Influences of Heredity and Musical Environment on the Scores of Kindergarten Children on the Seashore Measures of Musical Ability," Journal of Applied Psychology, XXIII (June, 1939), 348.

were administered the Seashore tests so Friend could compare the performance of parents and their child. For the most part, the correlations were positive but low. Correlations of scores between mother and child were slightly higher than those between father and child.[37]

Another music aptitude test, the MAP, has been used with younger children than it was originally designed to test. Harrington used the MAP with second- and third-grade students.[38] DeYarman adapted Harrington's version to measure the music aptitude of kindergarten and first-grade students.[39] DeYarman developed new kinds of answer sheets and modified test directions for these subjects. Both researchers used only three subtests of Gordon's original battery.

---

[37]Ibid., p. 355

[38]Charles J. Harrington, "An Investigation of the Experimental Version Primary Level Musical Aptitude Profile for Use with Second and Third Grade Students," Journal of Research in Music Education, XVII (Winter, 1969).

[39]Robert DeYarman, "Experimental Analysis of the Development of Rhythmic and Tonal Capabilities of Kindergarten and First Grade children" (unpublished Ph.D. dissertation, University of Iowa, 1971).

CHAPTER III

DEVELOPMENT OF THE RESEARCH TOOL

Preliminary research was conducted with three-year-olds and four-year-olds before the construction of the rhythm test was begun. This investigation was used to determine what types of testing procedures and test items the children of this age-group might successfully respond to. It was concluded from this work that the test would have to be administered individually because the children of this age could not write and attempts to test them in a group through individual verbal responses was not appropriate for gathering data. It was also concluded that testing periods would have to be short because of the relatively short attention span of the subjects of the age-group.

After trying various procedures, the author felt that he could not devise a test which would be appropriate for three-year-olds because most of them were not attentive to his attempts to test them for more than enough time to do one or two test items. Work was begun to design a test for four-year-olds based on the information and impressions of the preliminary investigation.

An underlying assumption was that the test should elicit both verbal and nonverbal responses. The logic behind this was that musical activities of children and adults involve some combination of both verbal and nonverbal behaviors. Verbal responses are defined as words such as "same" or "different" which are used to describe stimuli. The

nonverbal responses are those which involve clapping, tapping, or other bodily movement. It was arbitrarily determined that half the test items would require nonverbal responses.

The order of all four subtests and the order of items within the subtests were randomized. This was done in order not to superimpose any particular order on the test items. Each subtest has two practice items and ten test items. The test was recorded on reel-to-reel tape so that each administration would have identical stimuli. The following discussion concerns the procedures used to produce the test and collect reliability data on the test in a preliminary study with a group of 16 four-year-old subjects. Generally, the same procedures of test administration were used in the successive research.

One subtest was administered to every subject each afternoon. In the morning, the researcher worked with the subjects in groups of four to prepare them for the testing. The morning training will be referred to as "acclimatization." The subjects in each acclimatization group and the order in which each group was trained was randomized each day to compensate for the effects of interaction of the subjects, the disturbance of the subjects' daily routine, and the time of day.

Because many of the subjects may never have been exposed to a formal testing experience and had never met with the researcher, the acclimatization was used for the following purposes: to develop a rapport between the experimenter and the subjects; to acquaint the subjects with the tape recorder and other equipment; to acquaint the subjects with the procedure and types of responses required; and to have

each subject respond to stimuli which were similar to those on the subtests. In the afternoon, the experimenter briefly reviewed the morning acclimatization, gave the practice items, and then administered the rhythm subtest. The afternoon testing procedure was conducted with each subject individually. The complete procedure for each acclimatization is presented in Appendix A.

Today, most of the music listened to and performed by people living in this country has a rhythmic content which is based on a recurring pulse. Patterns of sound, often referred to as rhythms, are performed in relationship to the tempo or rate of these pulses. In constructing the rhythm subtests, the researcher takes the position that a child or adult who listens to or attempts to perform music with this type of rhythmic constitution responds with some degree of accuracy to the recurring pulses and patterns of sound within this type of music. This investigation is limited to these two rhythmic elements. These two elements will constitute the stimuli of the rhythm test, and a subject's accuracy of response to these stimuli will constitute his rhythmic ability.

As previously stated, it was desired to balance equally the required responses of the subjects to the stimuli between verbal and nonverbal responses. For this reason there are two subtests for the accuracy of response to pulse and two for the accuracy of response to rhythms (patterns of sound). In two of the subtests the subjects are tested for the ability to differentiate between like or unlike pulses and like or unlike rhythms. In the other two subtests the subjects are

asked to duplicate various rhythms and recurring pulses. It is conceivable that a child can tell that two series of pulses are the same or different, but the child may not be able to produce a steady series of pulses. Likewise, he may be able to determine whether or not two rhythms are alike, but he may not be able to reproduce a given rhythm. This test is designed to take into account these two factors.

The following is a more detailed discussion of the construction, procedures of administration, and scoring of each subtest.

Test A: Like-Unlike Tempi

The stimuli for this test were produced by a Seth Thomas metronome (wind-up model). The subject was presented with a pair of stimuli consisting of eight ticks each. The first stimulus (eight ticks) was at a set rate. Following a short pause, the second stimulus was presented at either the same or a different rate from the first. The subject was to indicate whether the stimuli were the "same" or "different." The tempi (rates) used in this subtest and in the other three subtests were the same. The ten tempi, which will be referred to as the "selected tempi," are as follows: 60, 66, 72, 80, 88, 96, 104, 112, 120, and 132 beats per minute. The researcher selected these tempi because they represent the middle range of tempi available on most metronomes and avoid extremely slow or fast tempi.

One each of the above tempi was selected for the first of each pair of the stimuli in Test A. The order of the tempi was randomized. Five tempi were randomly selected to be repeated unchanged. The other five were paired with a slower or faster tempo. Whether the second of

each pair would be slower or faster than the first was determined by random selection. Selection yielded three slower and two faster tempi. The second tempo of each pair of these stimuli was then set so it would be 24 beats per minute different from the first tempo. The researcher decided to use this deviation in rates because of findings of the preliminary experimentation. Generally, most children had much difficulty differentiating between tempi which were less than 24 beats per minute in difference. The two-second delay between presentations of the stimuli also seemed to complicate the subjects' decision process somewhat.

The items of Test A were graded as either right or wrong. Each item carried the numerical value of one point. The two practice items included one pair of stimuli with the same tempo and one pair with different tempi. Test A and the other three subtests are included in Appendix B.

Test B: Continue Taps

In this test the subject was presented with the sound of a ticking metronome set at one of the selected tempi. As soon as the subject began hearing the ticks he was supposed to imitate the ticks by striking together two small blocks. The metronome was silenced after it had given out with eight ticks; however, the subject was to continue tapping until he heard a bell ring.

The tape recording of this test was as follows. Two Seth Thomas metronomes were used for the stimuli. One was wind-up model; the other was an electrical model with a flashing light. The metronomes

were synchronized to produce the stimuli at the exact same instant. After eight ticks the wind-up metronome was silenced and the flashing light of the other metronome continued for eight more flashes. A small xylophone was struck simultaneously with the sixteenth flash of the light.

In the preliminary stages, the researcher had difficulty constructing and administering this subtest. At first, hand clapping was used for the children's responses. This was not satisfactory because some of the children clapped their hands too softly for the sound to be picked up by the tape recorder which was used to record responses. The researcher next tried two rhythm sticks (dowels) which were 29 centimeters long. Many children had problems holding the sticks and striking them together. Some children dropped the sticks quite often. After trying several objects, it was discovered that a pair of wooden blocks was the best vehicle for the tapping response. The square blocks measured four centimeters on each side, and they were two centimeters thick. The edges of the blocks were rounded. These blocks will be referred to as the "music blocks." It was found that the blocks tended to discourage distraction and play activities which often occurred when other items (triangles, drums, and claves) were used. The blocks made a sharp click when struck together, and the sound was recorded well by the tape machine.

Another problem was that the subjects would often stop the tapping before they were supposed to. In the preliminary research a blow on the claves was used to indicate the point in time when the sixteenth

beat had been reached. This was for the purpose of scoring the test. The researcher found that when the bell was substituted for the claves and the researcher made more of a game out of the test, then, the children responded more reliably and with greater motivation.

The subjects' responses were recorded on a small Hitachi tape recorder. Two doctoral students in music education at The University of North Carolina at Greensboro were used to judge the responses. The judges were asked to listen to the tape and count the number of taps produced by the subject between the eighth tick of the metronome and sounding of the bell. If a subject was rated by the judges as having completed eight taps, then the subject received one point for that item.

The judges were permitted to hear an individual item as many times as they wished. The judges scored 160 items in the preliminary study. They agreed on the scoring of all items except four. These items were given a point value represented by one-half point. The researcher had previously scored the items before the judges did. The correlation of the judges' and researcher's scorings was r=.987. The results indicated the test could be reliably scored by one judge. All scoring in successive research was done by the researcher alone.

Test C: Rhythm Duplication

In this test the subject was presented a short rhythm pattern and asked to duplicate it. The response was recorded and later graded by the same judges who scored Test B. First, the judges were asked to determine the total number of taps (number of times the music blocks

were struck together) in the subject's response to a stimulus. The judges were asked to disregard how well the subject's pattern matched the stimulus. If the subject's response had the exact number of taps as notes in the original stimulus he was awarded one-half point. There was no difference in scoring any of the items by the judges on this point. The judging of the researcher agreed. If the subject got the correct number of taps in an item, then, he could be awarded another one-half point on each item which generally matched the rhythm of the stimulus. The judges were told to: 1) award the half point if the response resembled the stimulus in the relationship of longer to shorter notes within the stimulus, 2) disregard any differences in tempo between the stimulus and the response, and 3) award no credit unless the relationship of longer to shorter notes was consistent within a particular response. The judges agreed on the scoring of 157 of 160 items. Concerning the items which the judges did not score alike, these items were all given a value of three-fourths of a point. The correlation between the judges' scoring and that of the researcher was r=.989. The results indicated that the test could be reliably scored by the writer alone, and all future administrations of this subtest were scored by the writer himself.

Data from test scores of subjects used in the preliminary research is based only on the scoring of the two judges. Subsequently collected data is based on the researcher's scoring only.

The ten rhythm patterns were constructed so that no slurred or tied notes were included. Each note sounded separately, and there were

no rests. There were pairs of three-, four-, five-, six-, and seven-note patterns used in this subtest. The researcher made the assumption that at least three notes were necessary to establish some sort of pattern, and he felt that patterns having more than seven notes would be too complex for most four-year-olds.

The tempi of the patterns were randomly selected from the selected tempi, and each pattern was performed on a different instrument when the tape was recorded. Four instruments of definite pitch were used: marimba (A below Middle-C), piano (Middle-C), electric piano (E above Middle-C), and xylophone (G above Middle-C). The patterns played on these instruments were performed on the single pitch as indicated above. These pitches were selected because they fall within the general singing range of four-year-olds. Pitches were randomly assigned.

The other six instruments were percussion instruments of indefinite pitch: tom-tom, rhythm sticks, snare drum, tambourine, cowbell, and claves (some authorities may consider the claves to be of definite pitch).

The two practice items were five-note patterns at the tempo of 93 beats per minute. This represents the mean of the selected tempi and the mean of the number of taps per pattern.

Test D: Like-Unlike Rhythms

In this test the subject responded to two short rhythms by saying "same" or "different." The test was similar to Test C in that the selection of instruments and tempi were identical. The instruments of definite pitch were randomly reassigned the pitches used in Test C.

None of the rhythmic patterns used in this subtest were identical with those used in the previous subtest. The number of notes per pattern ranges from three to seven as in Test C. Five patterns were selected to be repeated without change. The other five were modified on the repetition. The modified versions had the same number of notes as the unmodified ones, and both patterns were one measure in length with the same meter signature (or time signature). Whether the original version or the modified version was presented first was determined randomly. Test items were graded right or wrong.

Tape Recording the Test

The test was recorded using two identical Sony tape recorders (Model TC-105 A). The test was also administered using the same model of the Sony machine. The test was recorded in monaural sound at the speed of seven and one-half inches per second.

There are 30 inches of leader tape between the test items on all subtests except Test C which has 60. The tape was added so the tester would have enough time to turn the tape machine off before running into the next test item. Subtests A and D involve the subject's response to two stimuli. Leader tape 13 inches long was spliced between the end of one stimulus and the beginning of the next one. Through a process of cueing, the experimenter discovered the beginning and end of each of the various items within the test. Next, the researcher added one inch of recording tape to the end of the final sound of each stimulus and one inch before the beginning sound of each stimulus. The 13 inches of leader plus the two inches of recording tape created a two-second time

interval of silence between the stimuli of each item. The researcher believed that this was a reasonable length of time.

Subtests B and C involved the recording of the subject's responses. In Test B both the machine playing the stimulus and the machine recording the response were turned on simultaneously. Shortly after the bell rang both machines were shut off together. There were 60 inches of leader added after each stimulus in Test C, as was previously mentioned. This extra leader allowed each subject eight seconds for responding. If the subject did not begin to respond after a few seconds, the experimenter said "go ahead" in a soft voice. As soon as the subject began to respond the tape machine playing the stimulus was shut off, but the other machine ran until the subject finished. Each of the subjects was required to begin the response within a period of eight seconds.

The tape recording of the test was made by the writer and two assistants both of whom were majoring in music education. All of the subtests were recorded on one of the Sony machines. All of the test items of Test D were recorded on one machine and rerecorded on the master tape. There, of course, was a small drop in sound quality through this procedure; however, the drop in quality was the same for all items. The researcher felt that this was important to insure that all items which were supposed to be alike would be. This procedure eliminated any nuances of difference which could be caused by accent or rhythmic inaccuracy. Any differences in accent between the unlike patterns would be to the benefit of the subjects when deciding whether patterns were alike

or different.

All test items were recorded and rerecorded until all involved agreed that the items were musically correct and accurate according to the notation and tempi markings. This often involved recording a stimulus ten or more times.

Both metronomes were tested for accuracy before and after the recording sessions. On all occasions the metronomes were completely accurate or accurate within one tick of 60, 96, and 132 beats per minute when tested for one minute. The following is a discussion of the subjects used in the preliminary administration of the test and the events surrounding the period in which the test was administered.

The subjects who took part in the testing were children in a four-year-old group at Hester's Creative School on Spring Garden Street in Greensboro, North Carolina. There were 38 children in this group. The children were assigned to four groups according to their age as of August 1, 1972. The grouping was as follows: 49-51 months (group 1), 52-54 months (group 2), 55-57 months (group 3), and 58-60 months (group 4). Two children of each sex were selected at random from each group. The subjects totaled 16. Sex and age were accounted for through the selection process; however, this was a sample of subjects who were from a middle-class environment. The school is a tuition-operated private school. The director of the school, Mrs. Henrietta Hester Harris, told the researcher that the children in her school were predominantly from middle-class families.

Health records were available at the preschool for all the subjects. They were checked for any mental or physical impairments.

Health records revealed that none of the subjects appeared to have any condition which could cause a reasonable doubt that all subjects were normal both physically and mentally. After the researcher explained the purposes and exact procedures of the study, Mrs. Harris concurred with the researcher that all subjects should participate.

It was felt that each subject should understand and be able to verbalize the words "same" and "different" before they took the test. Understanding these terms was crucial to the performance on the subtests, especially Subtests A and D. Since this was the researcher's first encounter with the subjects, he was not sure that these terms were used and understood by these subjects. They had not been observed previously in any of the preliminary research. A test was devised to see if the subjects could use the terms. This test consisted of pairs of visual stimuli which were presented to the subjects. The subjects were to say if the pair of stimuli were the "same" or "different." This test was administered to the subjects one week prior to the administration of the first rhythm subtest.

The stimuli were pairs of shapes (triangles, squares, circles, rectangles, and other shapes) cut out of construction paper and pasted on the black pages of a photo album. The stimuli were of different colors, but both stimuli in each test item were the same color. This procedure was used in order to avoid complications which might be caused by a color-blind subject; however, the writer had no reason to believe any subject was color-blind.

There were twelve test items divided into two sections. Each section had six items. The first six items consisted of pairs of

shapes which were pasted side by side on the same sheet. Two pairs had the same shape; two pairs had the same shape, but one shape had ink markings on it; two pairs had different shapes. The second section of the test was like the first except only one shape appeared on each page. This tested the subjects' ability when the stimuli could not be directly compared.

Six practice items were given to each subject before the visual test. There was one practice item for each of the six possible combinations of stimuli. Whether a pair of shapes was to be a practice item or one of the test items was determined by chance.

The test was administered individually, and no acclimatization was used except for the practice items. The Kuder-Richardson reliability test was applied to the data which yielded the value .670. The mean for the correct number of responses to items on this test was 10.00 for the twelve items. Five subjects made a perfect score. One subject missed six items which was the most incorrect responses of any subject. The next day the researcher went over the results of the test with each child individually and pointed out both the correct and incorrect responses of each child. This was done for subjects who made a perfect score also. The writer wanted to insure that all subjects understood the terms "same" and "different" well. After going over the test, those subjects who made at least one error on the first test were retested. Testing was again individual. All subjects except one made perfect scores on this administration. A few hours later the review-test procedure was used with this subject. The subject did not miss any of the

items on the third administration. The researcher was then reasonably sure that the subjects could use the terms "same" and "different" accurately and reliably, and he believed that they were now ready to begin acclimatization for the rhythm test.

The researcher was permitted to use a room in the preschool for the investigation. The room was quiet, and the researcher and subjects were able to work undisturbed. The subjects took the subtests while seated three feet from the speaker of the tape recorder. Masking tape was placed on the floor to mark the position of the chair. The tester was sure that the chair was in the same position for each subject. In acclimatization the children sat in a semicircle in front of the tape recorder. During testing and acclimatization the researcher sat behind the tape recorder and faced the subject or subjects.

The researcher realized that there would be times during the test when a subject would not respond after a test item because he or she was not paying attention. This happened several times during the preliminary work with the test. The tester used the following methods of determining if the subject had not paid attention to the stimulus he was supposed to respond to: the subject did not respond immediately after a stimulus, even after being encouraged to respond by the tester; the subject talked during a stimulus. If the subject had not paid attention, in the opinion of the tester, the test item was repeated. Most subjects went straight through the subtests without interruptions. A few asked questions, talked, or made sounds during some of the testing. The researcher was careful to answer all questions and take time

with a subject who became distracted for one reason or another.

The four acclimatizations and rhythm subtests were administered to the subjects by the researcher between August 8-11, 1972 in the manner previously discussed. The writer computed split-half reliability for the test as a whole and for the separate subtests using the Pearson product-moment formula. The reliability for the test as a whole was r=.903. The reliabilities for subtests A, B, C, and D were respectively .17, .92, .95, and .65. The data from a small sample indicated that the test was probably reliable enough to be seriously considered for more investigation.

In order to measure the homogeneity of the test items in terms of how the individuals responded to them, the author applied to the data the Kuder-Richardson reliability coefficient as developed by Hoyt using analysis of variance. The data yielded the coefficient of .736. This statistic indicates that the test may possibly have fairly high validity; although, as of this time, the author had not studied the validity of his test. In a discussion of the Kuder-Richardson reliability coefficient, Bruning and Kintz state:

> A high reliability coefficient (.70 or higher) would mean that the test was accurately measuring some characteristic of the people taking it. Further, it would mean that the individual items on the test were producing similar patterns of responding in different people. Therefore, a high value would mean that the test items were homogeneous, therefore, valid.[1]

One more matter concerning the testing remains to be discussed. The researcher gave the subjects M&M candy at specific times during the

---

[1]James Bruning and B. L. Kintz, Computational Handbook of Statistics (Glenview, Illinois: Scott, Foresman and Company, 1968), p. 191.

experiment. Soon after the subjects came to the testing room for the acclimatization, each was given one piece of M&M candy. This was done before any work was done with the test or testing procedure. After the acclimatization was over, each child was given two pieces of the candy just before leaving the room. The same procedure was followed when the subjects were actually tested.

It was not the intention of the researcher to condition any particular behavior. The candy was given by the researcher to develop a rapport with the children. One might argue that the researcher was conditioning the subjects' behavior of coming to the acclimatization and testing situations. Also, the behavior of going through the acclimatization or taking the test may have been reinforced. If this candy motivated the subjects, then, it seems that the candy was a beneficial asset to the research.

One might also argue that giving candy after the test may have conditioned a certain type of response to the last test item. This is doubtful because there was an interval of at least a half-minute or longer between the administration of the last item and the giving of the candy. In this interval the researcher talked to the subjects. In almost all cases the researcher gave the candy when the child was smiling. This may have been reinforcing smiling behavior. The children did not seem to be aware that the last item was being administered when it was presented.

The candy was kept out of sight during the acclimatization and testing. The colors of candy and the order it was presented to the

children were randomized.

The previously described procedures of giving candy were also used in the experiment. At no time was candy given except during the actual testing and test acclimatizations.

CHAPTER IV

PROCEDURE

The four-year-old subjects at Hester's Creative School were given the picture test and the researcher's four rhythm subtests in August, 1972 in the manner and sequence discussed in the previous chapter. This preliminary research was done to collect data on the reliability of the test and to see how four-year-olds react to the testing procedures in general.

In order to collect data which could be used to explore the validity of the researcher's test, the researcher administered both his test and the MAP to 32 fourth-grade subjects. The fourth-grade subjects were used because this is normally the youngest age-group to which the MAP can be given. The subjects were randomly selected from 60 students in two classes at Chapel Square Elementary School in Annandale, Virginia. The subjects were randomly assigned to two different groups (Group I and Group II). Each group had eight boys and eight girls. The subjects of Group I were given Subtests A and D of the researcher's test on the morning of June 6, 1973. The acclimatization for these subtests were given in abbreviated form to all these subjects in a group. Also, the subjects took the subtests in a group using an answer sheet. In the afternoon the subjects were individually given the acclimatizations and Subtests B and C.

Group I and II were given the MAP according to the directions in the test manual on June 7, 8, and 11. The groups took the subtests

together in a large room, and the seating was randomized each day for each subject. One complete subtest (or section) was given on each morning of the previous dates. On June 13, 1973, Group II was given the researcher's rhythm test. The same procedure was followed that was used for Group I. The previously discussed testing schedule was arranged to compensate for any effects which taking the tests may have had on the performance of either test.

The picture test was not given to either group, and none of the subjects received candy during the testing procedure.

The researcher tried to raise the validity of his test through modifying the original test items which did not have an item difficulty of between about .20-.85 and an item discrimination of about .20 or greater. This was based only on data from the scores of the Greensboro subjects at Hester's school. Items in the following subtests were modified:

Subtest A
- 3. Stimuli were increased to 138 beats per minute
- 5. Switched in sequence to item number eight
- 8. Switched in sequence to item number five

Subtest B
- No changes in this subtest

Subtest C
- 21. Changed from a six to a five-tap item
- 24. Changed from a seven to a five-tap item
- 25. Changed from a five to a four-tap item

Subtest D
- 32. Rerecorded in original form using claves
- 33. Changed from a four to a six-tap item
- 40. Changed from a five to a seven-tap item

These modified items are located in Appendix B as items in parentheses.

These item changes constituted a new form of the rhythm test. The test was administered to 10 four-year-old subjects (five boys and five girls) in order to collect data on the new test items. The subjects were members of the Bel Pre Day Care Center in Wheaton, Maryland. The subjects were given the picture test on June 25, 1973, and they were given one subtest of the rhythm test on each day of the period of June 26-29. The same procedures of test acclimatization, individual testing, and giving of candy were used with this group as were used with the Greensboro subjects. The purpose of this phase of the research was to collect data to see if the changed items had improved in item discrimination and item difficulty. These data will be presented in the next chapter.

In a major part of the present research, the experimenter was seeking to determine what effect training would have on the performance of four-year-old preschool subjects on a posttest of the rhythm test. An experimental group and two control groups were used to determine the effects of the training. All three groups were given the new form of the test as a pretest. The experimental group then received 20 training sessions. One control group received no training and did not have any contact with the researcher until the posttest. The other control group met with the researcher and participated in nonmusical activities for 20 sessions which were equal in length to the training sessions for the experimental subjects. This was done to determine if personal contact with the researcher would effect posttest scores. Each group was retested after one month.

The 36 subjects who participated in this research were chosen from a pool of 62 four-year-old preschool students at five preschool centers in Montgomery County, Maryland. Each student was placed in one of three categories according to his or her age as of July 1, 1973. These categories represented the following ages: 49-52 months, 53-56 months, and 57-60 months. In the preliminary research with the Greensboro subjects the researcher had been able to divide the subjects into four groups according to which quarter of the year each was born; however, he was unable to do this with the Montgomery County students because of the unequal distribution of birth dates among the students.

Next, each age-group was subdivided according to sex. This made a total of six subgroups. Two subjects were randomly assigned from each subgroup to each of the following: an experimental group (Group X), a control group (Group C), and a control group which had contact with the researcher between the pretest and posttest (Group CC). This made a total of 12 subjects in each group. Six more subjects were randomly selected and assigned to each of the three groups to act as substitutes in case of attrition.

The subjects whose data were used in this research represent children from a fairly wide variety of economic, cultural, and racial backgrounds. The five preschool centers which the subjects attended are members of the Montgomery County Child Day Care Association. This association uses a sliding tuition scale which depends on family income. Some subjects came from very wealthy families; others were members of families partially supported by welfare assistance. There was also a

number of subjects which came from middle-income families.

The data of the 36 subjects used in this research have a racial-cultural breakdown as follows: 24 white, 5 black, 4 Spanish, 2 Indian, and 1 Korean. Race was not a factor in the assignment of subjects to the various groups.

The following is a list of the preschool centers and the number of subjects which participated in the research:

Kensington Day Care Center (8)
Del Ray Day Care Center (8)
Westmoreland Day Care - Nursery School (8)
River Road Child Day Care Center (6)
Bel Pre Day Care Learning Center (6)

The testing procedures used with these subjects were slightly modified from those used in preliminary research. The changes were necessary because of the enormous amount of travel time it required for the researcher to visit the five schools each day. In the pretest the subjects were first given the picture test. Soon after taking the picture test, each subject was informed of his performance on each item of the test. In preliminary research the experimenter had waited much longer before going over the picture test results. The same procedure was used for those subjects who had missed one or more test items as was used in preliminary research procedures. The picture test was not given in the posttest phase.

The modified test form was used for both the pretest and posttest. Only one subtest was given a subject each testing day; however, individual testing began soon after the group acclimatization at each center and was not put off until the afternoon. Further, some group

acclimatizations and testing was done in the afternoons. The order in which the researcher visited the various schools was randomized as much as possible within the framework of the daily schedules of the schools. For example, one school went on a field trip on Thursday mornings. One school had swimming on Friday mornings. All work with subjects in these preschools had to be done in the afternoons of the days when the subjects were busy in the mornings.

All subjects who participated in this research were free from any hearing defects. A hearing test was given to all subjects of the centers used in this research by nurses from the Montgomery County Child Day Care Association in May, 1973. Also, the researcher discussed each subject and substitute subject with that subject's teacher or teachers and the head of the preschool which he attended. One subject who had been randomly selected to participate had to be replaced in the opinion of the subject's teacher because the child had certain problems which the teacher felt should exclude this subject.

The substitute subjects were tested and received training, contact, or no contact along with the subjects in their group who had been selected to participate in the findings. As far as the researcher knows, no subject knew which group he was in or whether he was a subject or substitute subject. The attrition of the subjects was as follows: Group X (2), Group C (2), Group CC (1). The substitution of a subject was made by random selection of a substitute subject from the same age and sex category as the subject who dropped out because of vacation or leaving the center for an extended period of time. One

subject had to be replaced because she did not want to participate.

The subjects of Group X and Group CC were given the picture test on July 6, 1973. These same subjects were given the rhythm test between July 9-12, 1973. After a month the subjects were given the posttest between August 13-16. Group C was given the picture test on July 13, 1973. The rhythm pretest was given between July 16-19, and the posttest was given between August 20-23. None of the subjects in the three groups received group acclimatization before the posttest. Candy was given during the pretest and posttest phases in the same manner as previously discussed in preliminary research. No subject in any group received candy from the researcher between testings.

When the Group C subjects were tested they were one week older than the other subjects. This could not be avoided since time did not permit the testing of all three groups within the same week. It is probably not likely that the subjects in Group C had an advantage.

Between testings, each subject in Group CC met with the experimenter for about 12 minutes per day on Monday through Friday for a month (20 meetings). All subjects at each school met together as a group. The groups ranged in size from two to four children. During these sessions the researcher read stories and engaged the subjects in learning situations of a nonmusical nature. A list of the activities for each session is included in Appendix C. If a child missed one of the sessions he or she was allowed to make up that session. In this situation the subject was paired with another student who was in no way involved in the experiment and the researcher did the same thing that had been done in

the missed session.

Between testings, the subjects of Group X also met for the same number of minutes and sessions with the researcher. At each center the subjects of both Group X and Group CC used the same room but at different times. The order in which the groups met at each center was random with each group meeting first 10 of the 20 times. Also, the order in which the researcher visited the schools was randomized each day as much as possible within the framework of each schools' schedule.

The training for the subjects in Group X was as follows. At each center the subjects in this group were allowed to participate equally in the following:

1. Determining if two tempi produced by a metronome were the same or different
2. Reproducing on a musical instrument (percussion) various tempi produced by a metronome
3. Duplicating rhythm patterns produced by the researcher
4. Determining if two rhythm patterns produced by the researcher were the same or different

These activities represent the four abilities measured by the rhythm test. Each week during the period of training one of the above activities made up an entire training session. None of the activities used had stimuli which were exactly like any item in the researcher's test. This left one day of each week which had no training activity. On this day the subjects were engaged in one of the following: marching to a phonograph recording of band marches, clapping the rhythm of simple tunes appropriate for preschool children, walking in the rhythm of simple tunes, or clapping to band marches (two beats to the measure).

The week in which each of the previously discussed activities occurred was randomly determined. Each activity was included with the four training activities for a given week and the order of all five was randomized.

If a subject in Group X missed a training session he or she was allowed to make it up in the same manner used for Group CC. A detailed discussion of the training sessions, order of the sessions, and stimuli presented in the sessions is included in Appendix D.

The researcher conducted a final data collection to gain more information on the validity of the new form of the rhythm test. Ten kindergarten students from the J. Enos Ray Elementary School of the Prince George's County Public Schools in Maryland were the subjects. Five boys and five girls were randomly selected from the morning kindergarten class of 32 students. The subjects were given the picture test and the four subtests of the rhythm test on January 7-11, 1974. No candy was given any of these subjects at any time because of school regulations governing research.

The researcher had the subject's teacher and music teacher (who met with them once per week) rank the 10 subjects according to what they considered to be their musical rhythmic ability. The paired comparison technique was used for the ranking. The purpose of this piece of research was to collect data to be used to correlate scores on the rhythm test with teachers' ratings of rhythmic ability. The remainder of this chapter will be devoted to the methods used in treating the data.

Split-half reliability of the researcher's rhythm test was computed using the Pearson product-moment formula. Each correlation was corrected using the Spearman-Brown prophecy formula to approximate the reliability for the full length of the test. This procedure was used on data from preliminary research with the Greensboro subjects; the fourth-grade subjects; the Bel Pre subjects used to gain data on the item difficulty and discrimination of the modified items; the subjects from Montgomery County used in the experiment to determine the effects of training on posttest scores; and the kindergarten subjects from Prince George's County. This procedure was used for the test as a whole and for each subtest.

The test-retest reliability of the rhythm test was computed for the whole test and each subtest via the Pearson product-moment formula. This was done only with the data of the preschool subjects of Montgomery County participating in the experiment to determine the effects of training on posttest scores since these were the only subjects which took the test twice. The scores of all three groups were taken as a group. It should be noted that there was an intervening variable (the contact with the researcher or training) between pre- and posttest for the experimental group and control group which had contact with the researcher. The test-retest reliability was also computed for the subjects of Group C taken separately since this group was the only one which had no intervening variable between testings.

Scores on the rhythm test of fourth-grade subjects were correlated with their scores on Gordon's MAP. The MAP was administered and

scored according to directions in the test manual. The MAP yields the following scores: composite score for the complete test, total for the "Tonal Imagery" section, "Tonal Imagery Part I" (melody), "Tonal Imagery Part II" (harmony), total for the "Rhythm Imagery" section, "Rhythm Imagery Part I" (tempo), "Rhythm Imagery Part II" (meter), total for the "Musical Sensitivity" section, "Musical Sensitivity Part I" (phrasing), "Musical Sensitivity Part II" (balance), "Musical Sensitivity Part III" (style). All the above scores were correlated with the four subtests and the composite score of the researcher's rhythm test. The Pearson product-moment formula was used.

The Spearman rank-order formula was used to correlate the scores received on the rhythm test by the kindergarten subjects of J. Enos Ray Elementary School with their music teacher's and classroom teacher's rankings of rhythmic ability. As was stated previously, the rankings were derived using the paired comparison technique.

The intercorrelations of subtests of the researcher's test were made with the Pearson product-moment procedure. This was performed on data from the experiment concerning the effects of training. This was done to determine the degree to which the subtests correlated perhaps giving some information as to whether the subtests seemed to be measuring the same or different traits of rhythmic ability.

As was stated before, each test item was investigated to determine its item difficulty and item discrimination. This procedure was applied to data from the subjects having the one-third lowest and one-third highest scores in the testing with the Greensboro subjects. The

acceptable range of item difficulty was set around .20-.85; the acceptable range of item discrimination was about .20 or greater. The items which did not fall generally within these ranges were modified. The new form of the test was administered to 10 four-year-old subjects before the test was used in the experiment on the effects of training. This procedure was used to raise the validity of the researcher's test. The following is a discussion of the treatment of data collected using the modified test form.

The performances of the three Montgomery County groups of subjects in the training experiment was evaluated in terms of their performance on the various items in the four subtests. The evaluations were made on the performance of all subjects, taken as a single group, on the items in the pretest. This was done to avoid any effects which may have been brought about by training or contact with the experimenter.

The data was examined to determine if subjects did significantly better on the items which have different stimuli in Subtest A when a slower tempo is followed by a faster tempo. This was computed using chi-square. Significance was set at the .05 level or beyond.

In Subtest A the Pearson product-moment formula was used to find the correlation between the number of correct responses to an item and the tempo of the stimuli in items which use the same tempo. This was done to discover information on the rate of stimuli and the ability of the four-year-olds to discriminate correctly when stimuli are alike.

In Subtest A the correlation was made between the correct number of responses to items having different stimuli and the midpoint of the

tempi of these items. This was done to uncover information about the range of different tempi and a four-year-old's ability to discriminate a difference.

Again, using the Pearson product-moment formula, the correlation was made between the number of correct responses to items in Subtest B and the tempo of the items. This may give some evidence about whether a subject of this age-range can more accurately duplicate a slower, medium, or faster tempo-range.

In Subtest C the researcher computed the correlation between the number of taps in an item and the mean number of correct tap-responses to items having that many taps. This may indicate whether or not the more taps there are in an item the more difficult it is for a subject to duplicate that number of taps correctly.

In Subtest D the experimenter computed the correlation between the number of taps in an item and the mean number of correct responses to items having that many taps. This may indicate whether or not the more taps the stimuli of an item have the more difficult it is for the subjects to discriminate whether the stimuli are the same or not.

Hypothesis 1, that training will significantly improve posttest scores, was tested using analysis of covariance. In this study a significant difference exists among the performances of the three groups' posttests if an F ratio was significant at or beyond the .05 level. Each separate use of analysis of covariance in the treatment of the data in this study was preceded by an F-maximum test to determine the homogeneity of variances of pretest scores.

Hypothesis 2 was tested using analysis of covariance with the data of pre- and posttest scores of the two control groups. The level of significance of .05 was set to be required to discount the validity of the hypothesis assuming that Group CC had a greater mean posttest score than did Group C. An F ratio was computed for the composite scores as well as for the scores of each individual subtest; however, only the F ratio from the composite scores was used to validate or invalidate the hypothesis.

Hypothesis 3 was tested using analysis of variance. To reject the hypothesis that sex does not significantly affect the performance on the pretest of the researcher's test, an F value must be at or beyond the .05 level to be significant.

The researcher sought to determine if there was a significant improvement on the posttest scores for each of the three groups of subjects participating in the experiment. Each group was treated separately. A t value was calculated using a t-test. The value of the t would have to be large enough to be significant at or beyond .05. This data will not affect the acceptance or rejection of Hypothesis 1.

Each age-group's data was subjected to an analysis of covariance in order to determine whether there was a significant difference in performance on the posttest among the control groups and the experimental group within each of the three age-ranges. This data will not directly bear on the acceptance or rejection of Hypothesis 1. The test as a whole and each individual subtest was subjected to the analysis of covariance.

The researcher correlated the age in months of the subjects who participated in the experiment with their pretest scores for each of the subtests and for the test as a whole. The Pearson product-moment formula was used.

Finally, an analysis of variance was made on the test data of the Greensboro subjects in the preliminary research with the test scores of the fourth-grade subjects in Virginia. This was done to find out whether older subjects (fourth-graders) tended to perform significantly better on the old form of the test than did four-year-olds. The level of significance was set at .05. The same procedure was used to determine any significant differences between the performance of the four-year-old subjects in the experiment (using pretest scores) and the kindergarten-age subjects. Scores were based on performance on the new form of the test for these two groups.

The data including test ranges and test means are presented in the following chapter.

CHAPTER V

EVALUATION OF THE DATA

Table I below contains data on the test means and ranges for subjects who took the researcher's test but who were not included in the effects of training experiment. The table also includes information about the form of the test taken, age-group of the various subjects, and the number of subjects in each group.

TABLE I

MEANS AND RANGES OF THE RHYTHM TEST SCORES FOR SUBJECTS WHO DID NOT PARTICIPATE IN THE STUDY OF THE EFFECTS OF TRAINING ON POSTTEST SCORES

Source of Subjects	Number of Subjects	Age Group	Form of Test	Test Range	Test Mean
Hester's Center	16	4-yr.-olds	old	19.2	18.3
Bel Pre Day Care	10	4-yr.-olds	new	21.0	19.5
Prince George's Co.	10	kindergarten	new	18.0	19.7
Chapel Square Elem.	32	fourth-graders	old	16.0	31.7

Table II, which follows on the next page, includes test data of those subjects who did participate in the experiment. This table has data for all three subgroups taken as one group and for each separately.

TABLE II

MEANS AND RANGES OF THE RHYTHM TEST SCORES FOR THE MONTGOMERY COUNTY FOUR-YEAR-OLD SUBJECTS WHO PARTICIPATED IN THE STUDY OF THE EFFECTS OF TRAINING ON POSTTEST SCORES

Source of Subjects	Number of Subjects	Form of Test	Pretest Range	Posttest Range	Pretest Mean	Posttest Mean	Increase in Mean
All as a Group	36	new	24.5	22.5	17.7	20.4	2.73
Group X	12	new	21.0	20.5	18.9	22.6	3.74
Group CC	12	new	16.0	18.5	16.6	18.5	1.92
Group C	12	new	16.0	15.5	17.7	20.2	2.46

Note: Group X received training. Group CC had contact with the researcher between testings but received no musical training. Group C received neither training nor contact with the researcher between testings.

Since these subjects took the test twice, both pretest and posttest data are included as well as the increase in mean from pretest to posttest. Because two forms of the test were used in testing the various groups, it is not possible to directly compare the performance of the groups which did not take the same test form.

From Table I it can be seen that the four-year-olds who took the old form of the test had a test mean (18.3) which represents a score of slightly less than a 50 percent correct performance on the individual items of the forty-point test. The fourth-graders' mean (31.7) represents a slightly higher performance than 75 percent correct on the same items. The subjects who took the new test form were either four-year-olds or kindergarten-age. The means for these subjects lie close to a figure representing 50 percent correct performance on the items of the new test when the pretest scores of the subjects in the experiment are taken as a single group.

All three groups in the experiment had a mean increase from the pretest to the posttest. The group which received training had the highest increase; however, the group which had no contact with the researcher had a larger mean increase than the group which did. This data contradicts what the researcher had predicted would happen.

Table III includes data on the split-half reliability of the scores of subjects who were not in the experiment. Table IV contains similar data for pretest and posttest scores of subjects who were in the experiment. The reliabilities in these tables were corrected for the full length of the tests and subtests using the Spearman-Brown

TABLE III

SPLIT-HALF RELIABILITIES FOR THE SCORES OF SUBJECTS WHO TOOK THE RHYTHM TEST BUT WHO DID NOT PARTICIPATE IN THE EFFECTS OF TRAINING EXPERIMENT

Source of Subjects	Number of Subjects	Form of Test	Subtest A	Subtest B	Subtest C	Subtest D	Composite Total
Hester's Center	16	old	.17*	.79	.94	.65	.91
Bel Pre Day Care	10	new	.19*	.81	.90	.42	.89
Prince George's Co.	10	new	.84	.89	.91	-.74	.82
Chapel Square Elem.	32	old	.49	.30*	.12*	.42	.55

*The Pearson product-moment correlation to which was applied the Spearman-Brown prophecy formula was not at or beyond the .05 level of significance.

TABLE IV

SPLIT-HALF RELIABILITIES FOR THE SCORES OF SUBJECTS WHO TOOK THE RHYTHM TEST AND PARTICIPATED IN THE EFFECTS OF TRAINING EXPERIMENT

Source of Subjects	Number of Subjects	Form of Test	Subtest A	Subtest B	Subtest C	Subtest D	Composite Total
			P R E T E S T				
All as a Group	36	new	.44	.70	.90	.13*	.86
Group X	12	new	.39*	.67	.81	.61*	.86
Group CC	12	new	.30*	.86	.95	.43*	.91
Group C	12	new	.70	.53*	.95	-.61	.84
			P O S T T E S T				
All as a Group	36	new	.72	.81	.82	.31	.86
Group X	12	new	.66	.91	.78	.41*	.90
Group CC	12	new	.90	.70	.84	.59*	.88
Group C	12	new	.60*	.52*	.92	.53*	.70

*The Pearson product-moment correlation to which was applied the Spearman-Brown prophecy formula was not at or beyond the .05 level of significance.

prophecy formula. The split-half reliabilities of the preschool and kindergarten subjects are in a range of .82-.91 except for one which is .70. The scores of the preschool and kindergarten subjects seem fairly reliable when factors such as the age and maturity of these subjects are considered. The consistent measurement of the research tool is demonstrated by the fact that the pre- and posttest reliabilities for the subjects taken together in Table IV are the same (.86).

The reliabilities of the composite score totals for the four-year-old and kindergarten subjects (Tables III and IV) who took the new test, range from .82-.89 when the scores of subjects in Table IV are taken as a group-performance. The four-year-olds who took the old test have a reliability of .91 for the composite scores; however, the scores of fourth-graders who took the same test have a much lower reliability (.55). This lower reliability might be accounted for partially by the higher mean score and relatively small range of scores for the 32 subjects in this older and larger group. This might indicate that the test was too easy for the fourth-graders, and, consequently, it did not discriminate well.

The split-half reliabilities of the four subtests in both the old and new forms of the test is another matter. Excluding the subjects who were fourth-graders and using data from both the pretest and posttest for the three groups in the experiment whose scores are taken as a group, all four-year-old and kindergarten subjects who took the test had highest split-half reliabilities for the two subtests which required nonverbal responses (Subtests B and C). This may indicate the subjects

in this age-range more reliably respond to test items which involve nonverbal responses. The fourth-graders, on the other hand, had highest split-half reliabilities for the two subtests which were responded to verbally (Subtest A and D). There are probably many factors contributing to the more reliable performances by the younger subjects on Subtests B and C. One factor might be more motivation, which was caused by responding to something rather than uttering "same" or "different." Other factors such as a lack of "test-taking" sophistication may have affected the performance of the younger subjects. It was noted that for the younger subjects there was a definite pattern of better reliability of performance on the subtests involving nonverbal responses as defined in this paper.

Table V, which follows on the next page, contains data on the test-retest reliability of the scores of those subjects who participated in the experiment. There was an intervening variable between the pretest and posttest for two of the three groups. Only one of the correlations for the three subgroups taken together was a significant correlation at the .05 level. All correlations of Group C, which had only one-third the subjects, were significant. The .75 correlation for the composite total of Group C is not at the .85 level, but the figure does appear to be fairly high for the small number of subjects in this group. It is impossible to make a realistic statement about the reliability for the subgroups taken together because of the intervening variables and the insignificance of most of the correlations. The test-retest reliabilities for the subtests of Group C range from .29-.83. The

TABLE V

TEST-RETEST RELIABILITIES FOR THE SCORES OF SUBJECTS WHO TOOK THE RHYTHM TEST AND WHO DID PARTICIPATE IN THE EFFECTS OF TRAINING EXPERIMENT

Source of Subjects	Number of Subjects	Form of Test	Subtest A	Subtest B	Subtest C	Subtest D	Composite Total
All as a group**	36	new	.07*	.10*	.80	.49*	.45*
Group C	12	new	.34	.54	.83	.29	.75

*The Pearson product-moment correlation was not at or beyond the .05 level of significance.
**There was an intervening variable (training or contact with the researcher) between testings for two of the three subgroups.

correlation of Subtest C is the highest for Group C taken separately and for all subjects taken as a group. In Group C the next highest reliability is that of Subtest B which seems to indicate that nonverbal response to test items causes better reliability of performance. This is not upheld by the consideration that the second best reliability score on the subtests for all subjects as a group was that of Subtest D; however, this correlation was not a significant one.

Table VI contains the correlations in fourth-graders' scores between the composite and subtest raw scores of the researcher's test and the standard scores which the MAP yields. These data are based on the old form of the experimenter's test.

In terms of the highest correlations and significances of correlation, the researcher's test correlates highest with the following: the two MAP rhythm subtests (.60 and .61), the MAP composite rhythm score (.65), and the composite total for the MAP (.51). These correlations are moderate to moderately high; however, there are no minimum criteria which are universally accepted by researchers for the demonstration of the validity of a test using this particular method of validation.

The degree of correlation between the four composite scores of the MAP and the composite scores of the researcher's test seem to present a logical pattern for a valid rhythm test when correlated against a standardized musical aptitude test. The correlations with MAP rhythm subtests are highest, the correlations with the MAP composite total is next, and correlations with "Musical Sensitivity" and "Musical Imagery"

TABLE VI

CORRELATION OF RHYTHM TEST SCORES OF THE THIRTY-TWO FOURTH-GRADE SUBJECTS WITH THEIR SCORES ON THE MAP

Scores on MAP	Subtest A	Subtest B	Subtest C	Subtest D	Composite Total
Tonal Imagery I (Melody)	-.10*	.11*	-.28*	.18*	.25*
Tonal Imagery II (Harmony)	-.03*	.12*	-.26*	.17*	.23*
Composite Tonal Imagery	-.07*	.11*	-.29*	.19*	.25*
Rhythm Imagery I (Tempo)	.20*	.33	.14*	.44	.60
Rhythm Imagery II (Meter)	.21*	.34	.13*	.43	.61
Composite Rhythm Imagery	.21*	.36	.14*	.47	.65
Musical Sensitivity I (Phrasing)	.34	.16*	.32	.34	.47
Musical Sensitivity II (Balance)	.43	.14*	-.12*	.45	.47
Musical Sensitivity III (Style)	.23*	-.04*	-.12*	.26*	.29*
Composite Musical Sensitivity	.41	.05*	.01*	.45	.45
Composite Total MAP	.17*	.25*	-.09*	.41	.51

*The Pearson product-moment correlation was not at or beyond the .05 level of significance.

are lowest.

This pattern is also fairly consistent when the subtest scores from the researcher's test are correlated with the subtest scores and composite scores of the MAP. Half of the "Tonal Imagery" scores correlate negatively with the researcher's subtest scores. None of the positive correlations are above .19. None of the correlations with "Musical Sensitivity" is above .45, and one-fourth of the correlations are negative. Results which support the idea that the rhythm test is measuring rhythmic aptitude or ability are the facts that none of the correlations with "Rhythm Imagery" is negative and half of these correlations are above .30. However, many of the correlations in Table VI are not at or beyond the .05 level of significance.

The scores received by 10 kindergarten subjects on the new form of the rhythm test were correlated with their classroom teacher's rankings and music teacher's rankings of their "musical rhythmic ability." The paired comparison technique was used for the ranking, and the Spearman rank-order formula was used to compute the correlations which follow below in Table VII. The music teacher's rankings correlated higher, generally, than those of the classroom teacher, which is a reasonable expectation. The correlations of rankings and composite scores are positive, and the music teacher's ranking correlates fairly high (.78) with the composite scores. Also, the pattern of the correlations between the rankings and the subtest scores is fairly consistent between the teachers.

The correlation between the teachers' rankings computed with the Kendall rank-order method was $\underline{tau}$=.65. The correlation was significant

at the .05 level.

TABLE VII

CORRELATION OF RHYTHM TEST SCORES OF TEN KINDERGARTEN SUBJECTS WITH THEIR TEACHERS' RATINGS OF RHYTHMIC ABILITY

Teacher Doing Ranking	Subtest A	Subtest B	Subtest C	Subtest D	Composite Total
Music Teacher	.45	.88	.52	-.02	.78
Classroom Teacher	.55	.79	.48	-.25	.66

It is difficult to arrive at a completely satisfiable reason for the two low-negative correlations in the above table. A negative correlation was also found for the scores of one of the three groups of subjects in the experiment when split-half reliabilities were figured for that group (Table IV). Since both the rankings and score correlations for Subtest D in Table VII are negative, this represents a situation which might indicate that these results may have been caused by the correct guessing of some of the subjects.

Table VIII includes the intercorrelations of the subtest scores for the subjects who participated in the experiment. This data is from performances on the new form of the test. It is difficult to interpret the correlations since half of them are not significant at the .05 level. The mean of the correlations for the pretest is around .26 and about .30 for the posttest. Subtests A and B correlate quite low.

TABLE VIII

INTERCORRELATIONS OF RHYTHM SUBTEST SCORES FOR THE SUBJECTS WHO PARTICIPATED IN THE EFFECTS OF TRAINING EXPERIMENT

Test	A&B	A&C	A&D	B&C	B&D	C&D
Pretest	.07*	.45	.17*	.31	.39	.20*
Posttest	.09*	.50	.22*	.26*	.35	.36

*The Pearson product-moment correlation was not at or beyond the .05 level of significance.

This might indicate that the ability to decide whether two tempi are alike or different is not closely related to a subject's ability to reproduce a given tempo; however, the subtest correlations of Subtest A and B are not significant. The correlations between Subtests A and C, which are the highest of all the correlations, indicate that there may be some sort of relationship between a subject's ability to differentiate between like and unlike tempi and his ability to reproduce a rhythm pattern. The correlations between the subtests do not appear to be particularly high except for the two subtests A and C. This indicates that the subtests tend to measure different subtraits of rhythmic ability.

Table IX contains information on item difficulty and item discrimination based on data from preliminary research with the old form of the test. The items which are accompanied by an additional number

TABLE IX

ITEM DIFFICULTY AND ITEM DISCRIMINATION FOR THE ORIGINAL TEST FORM BASED ON DATA FROM THE GREENSBORO SUBJECTS ALONG WITH DATA ON THE REVISED TEST ITEMS FROM TESTING WITH THE BEL PRE SUBJECTS

Number of Item	Item Difficulty	Item Discrimination
Subtest A		
1	.25	.40
2	.62	.60
3	.81 (.75)	-.40 (.50)
4	.37	.20
5	.43 (.25)	-.20 (.00)
6	.56	.40
7	.62	.80
8	.37 (.50)	-.40 (.50)
9	.75	.20
10	.75	.40
Subtest B		
11	.25	.00
12	.43	.40
13	.25	.60
14	.12	.40
15	.37	.60
16	.19	.60
17	.43	.40
18	.31	.40
19	.31	.40
20	.37	.40
Subtest C		
21	.07 (.50)	.05 (1.00)
22	.31	.60
23	.19	.50
24	.00 (.12)	.00 (.50)
25	.15 (.62)	.30 (.50)
26	.34	.70
27	.34	.90
28	.21	.70
29	.03	.10
30	.31	.80

TABLE IX (Continued)

Number of Item	Item Difficulty	Item Discrimination
Subtest D		
31	.62	.20
32	.19 (.50)	-.20 (.50)
33	.81 (.62)	.00 (1.00)
34	.62	.20
35	.19	.20
36	.75	.20
37	.50	.20
38	.19	.20
39	.37	.20
40	.81 (.87)	.00 (.50)

Note: Items not in parenthesis are based on data from preliminary research with 16 subjects in Greensboro, North Carolina. The items in parenthesis are based on data in later research on a new form of the test item with 10 subjects at Bel Pre Day Care.

in parentheses are items which were modified to constitute a new form of the test. The new form of the test was given to 10 four-year-old subjects to collect data on the item difficulty and discrimination for the modified items. The numbers in parentheses are data from this subsequant research.

As was stated previously, the generally acceptable level of item difficulty was set by the researcher at .20-.85 and the level for discrimination was set at about .20 or greater. All the modified items improved in both discrimination and difficulty or improved on one of the two while not driving the other below the acceptable range. Items which had item difficulty of .19 were considered close enough to the set range to be acceptable. The researcher decided not to modify any of the items of Subtest B because of the balance between item difficulty and discrimination in this subtest. Item number 29 of Subtest C

was not modified because the researcher felt it was a very difficult item which might challenge the subject with a very high degree of ability. This item is one which involves the process of reproducing a rhythm pattern where guessing or luck are less strongly involved. The following is a discussion of data which is from the new form of the test.

The scores of the three groups participating in the experiment were evaluated in terms of the subjects' performances on the various items in the four subtests. These evaluations were based on the pretest scores for all the subjects taken as one group. This was done to avoid any complications which might have been brought about by the intervening variables.

The experimenter sought to determine if subjects did significantly better on the items of Subtest A which have different stimuli, when a slower tempo is followed by a faster one. This was computed using chi-square. The subjects did do significantly better (got more items correct) on the items having a slower rate followed by a faster rate. The chi-square value was significant at the .005 level.

In Subtest A, the correlation was made between the number of correct responses to an item and the tempo of the stimuli used in the item for those which have the same tempo for both stimuli. The correlation was -.42 and the significance of the correlation was at or beyond .001. This seems to indicate that the faster the tempo of items having identical stimuli, the more subjects tend to believe that the stimuli are different.

In Subtest A the correlation was made between the correct number of responses to items having different stimuli and the midpoint of the tempi of these items. The correlation was .37, but it was not significant at the .05 level. This tends to indicate that the faster the rate of stimuli in items having different stimuli, the more easily a subject can discriminate that the item has different stimuli.

Again, using the Pearson product-moment correlation, the correlation was computed between the number of correct responses to Subtest B items and the tempi of these items. The correlation was .77 and the significance of the correlation was .004. This indicates that, generally, the faster a stimulus between 60-132 beats per minute, the easier it is for the subject to reproduce that stimulus accurately. This finding somewhat corresponds to the finding of Williams which was previously discussed. Williams found that subjects were most successful in tapping synchronously with a stimulus rate which was between 80-100 beats per minute.

In Subtest C the researcher computed the correlation between the number of taps in an item and the mean number of the correct tap-responses to items having that many taps. The correlation was -.94 and the significance of the correlation was .009. This strongly suggests that when a subject is asked to duplicate rhythm patterns ranging from three to seven taps, his success in getting the correct number of taps in an item increases as the number of taps in the pattern decreases.

In Subtest D the correlation was made between the number of taps in an item and the mean number of correct responses to items having that

number of taps. The correlation was .66, but the significance of the correlation was .11. This may indicate, to some degree, that when the number of taps in patterns ranging from three to seven taps increase, subjects can better discriminate whether or not the patterns are alike.

The researcher tested Hypothesis 1, that training would result in significant improvement of posttest scores, using analysis of covariance. Separate analyses were made for each subtest and for the composite scores of the three groups of subjects in the experiment. The analyses were computed using posttest scores, and the covariates were the scores from the rhythm pretest. An F-maximum test was performed on the pretest scores to determine the homogeneity of variances. None of the variances was found to be significantly different at the .05 level. None of the F ratios generated through the analysis of covariance was sufficient to reject the null hypothesis. The null hypothesis was accepted.

Hypothesis 2 stated that the scores of the control subjects having contact with the researcher between testings would show a larger mean increase on the posttest than the scores of subjects not having the contact, but the larger increase would not be a significant one. From the data in Table II, it can be seen that the findings of this study are inconsistent with the hypothesis because the group which did not have contact with the researcher between testings (Group C) actually had the larger mean increase. The null hypothesis was accepted.

The investigator applied the analysis of covariance to scores of both groups for each subtest and for the composite test scores to see

if the improvement in the scores of Group C was significant. The analyses were computed using posttest scores, and the covariates were the rhythm pretest scores. An F-maximum test was applied to the pretest composite and subtest scores of both groups. None of the variances was found to be significantly different at the .05 level. None of the F ratios from the analysis of covariance was sufficient to be significant at .05.

Hypothesis 3, that sex would not significantly affect performance on the researcher's test, was tested by applying analysis of variance to the pretest data of boys and girls who took part in the experiment. The F value was not significant at the .05 level, and the null hypothesis was rejected. Neither sex seemed to perform significantly better than the other on the test.

The researcher subjected the data from the pretest and posttest scores of each separate group participating in the experiment to a t-test. The analysis of the data indicates that the posttest scores of Group X improved significantly over pretest scores. The level of significance was beyond .01. Group CC also improved significantly beyond the .05 level. Group C improved, but the improvement was not significant at the .05 level. These data indicate that the practice of taking the pretest may have helped the subjects do better on the posttest. In the case of Group X and CC, the training or contact with the researcher may also have contributed, in some way, to the significantly better performances by these groups on the posttest.

The pretest and posttest scores of each separate age-group's data were subjected to analysis of covariance to determine if there was

a significant difference in performance on the posttest among the three age-subgroups within the three groups. Separate analyses were made for the composite scores and each of the separate subtest scores for each of the age-subgroups. The analyses were computed using the posttest scores, and the covariates were the pretest scores. An F-maximum test was computed for each set of pretest scores used in each individual analysis of covariance. All of the variances were homogeneous except one which represented the pretest composite scores of the youngest of the three age-groups. None of the F values from the analysis of covariance was found to be significant. Most of the values were at a level of confidence which was greater than .20. Only one of the 15 F values was at the .10 level. This data indicates that the effect or lack of effect of training or contact with the researcher between testings is not a function of age among four-year-olds.

The researcher correlated the age in months of the subjects participating in the experiment and their pretest scores for each of the subtests and for the test as a whole. The data follows in Table X. The data seem to indicate that there may be some sort of moderate relationship between a four-year-old subject's age and his rhythmic ability as measured by the researcher's test. This supports the findings of Williams; Christianson; Jersild and Bienstock; and Vance and Grandprey which were discussed in Chapter II. These researchers concluded that, in younger children, rhythmic ability seems to increase with age to a degree and to a certain point in time. As was previously discussed, Vance and Grandprey got correlations as high as .37 for age and rhythmic

TABLE X

CORRELATION OF AGE IN MONTHS WITH SCORES ON THE PRETEST FOR SUBJECTS PARTICIPATING IN THE EXPERIMENT

Subtest A	Subtest B	Subtest C	Subtest D	Composite Total
.17*	.37	.30	.45	.46

*The Pearson product-moment correlation was not at or beyond the .05 level of significance.

ability in subjects below kindergarten-age.

The correlation of age with the performance on Subtest A is the lowest of the correlations and it is not a significant correlation. This might be so because the reliabilities for this subtest are rather low. It probably does not mean that the ability to decide whether two tempi-rates are alike is not related to age in months among four-year-olds. The low correlation could mean that the test is too difficult for this age-group.

Finally, an analysis of variance was made using test data of the four-year-old Greensboro subjects and the fourth-grade subjects at Chapel Square Elementary School. The F ratio derived was significant beyond the .001 level. This indicates that the older subjects perform significantly better on the researcher's test. The old test form was given to both of these groups. It could be that rhythmic ability is in flux for four-year-olds. Still, it could also be that four-year-olds

lack a certain amount of "test-taking" sophistication due to lack of maturity or other factors. The reason for this difference in performance seems unclear.

An analysis of variance was also made with data of the pretest from the subjects in the experiment and the scores of the kindergarten subjects. The new form of the test was used. The F value was not large enough to be significant at the .05 level. Apparently, the four-year-old and kindergarten subjects do not significantly differ in their performance on the researcher's rhythm test.

CHAPTER VI

SUMMARY AND CONCLUSIONS

Introduction

Many authorities believe that what happens to a child in his younger years can greatly affect the child's growth and development. At present, knowledge from research into the development of young children is incomplete and insufficient. This includes knowledge from research into the musical responses of preschool children. Two of the most neglected areas are the rhythmic response and the development of musical aptitude of preschool children.

The origin of music aptitude is not completely understood. Some authorities take the position that heredity plays the dominant role in the development or causation of music aptitude. Others feel that environmental influences contribute more. Still, there are those who feel that heredity and environment interact in some manner to give rise to music aptitude.

Two researchers who have worked in the area of music aptitude, Carl Seashore and Edwin Gordon, believe music aptitude is established somewhere around the age of ten; however, they do not seem to agree exactly on the effect of environment before this age-level. They do agree that after this age-level, aptitude does not seem to be strongly affected by training or experience. Some would disagree with this. Many researchers who do believe that environment does, perhaps, play an

important part in the growth of music aptitude also believe that the environmental influences on the young child are quite important to that growth. This is the position taken by the writer; however, it remains to be demonstrated how, if at all, environment can affect the growth of musical aptitude.

There are many unsolved questions related to the musical aptitude development in young children such as: Does it occur in spurts? Is it a gradual process? What experiences help a child reach a maximum music aptitude? What experiences thwart the growth of music aptitude?

In order for researchers to begin working on these problems some basic research tools are needed. The bulk of such tools employed to measure music aptitude have been designed for subjects in the upper elementary-level through adulthood. These are mostly the paper-pencil tests which are not suitable for the majority of preschool children who cannot read and write with enough facility to take these tests. Also, there are factors such as immaturity or a certain lack of "test-taking" sophistication which enter into consideration.

The purpose of this study was to develop a test designed to measure the rhythmic ability of four-year-old preschool children and to study the effects of training on posttest scores of the age-group. The term "ability" is used instead of "aptitude." This term was selected because it seemed to more appropriately describe a state of growth in young children when music aptitude had not been firmly established but was in flux. The rhythm test was designed for four-year-olds because they were the youngest age-group the researcher felt he could test

accurately after some preliminary investigation with three- and four-year-olds.

The study has two phases. One is data collection to determine the reliability and validity of the rhythm test. This phase also includes the modification of some of the test items in this forty-item test and research with the resulting modified test. The other phase is data collection to measure the effects of training on posttest scores using the modified test version. The following three hypotheses were investigated:

1. The test scores of four-year-old experimental subjects will significantly improve after training.

2. The scores of the control subjects having contact with the researcher between testings will have a larger mean increase on the posttest than the subjects not experiencing the contact; however, the larger increase will not be a significant one.

3. Among the subjects in the experiment, there will be no significant difference in the performance on the test which can be attributed to sex.

The Research Tool

The test is composed of four ten-item subtests which were played on percussion instruments and recorded on reel-to-reel tape. Subtest A involves a subject's determining whether two tempi-rates produced by a metronome are the same or different. In Subtest B the subject hears a metronome and tries to reproduce the ticking-rate by hitting two small wooden blocks together. The subject's accuracy in reproducing the rate after the stimulus is stopped is then measured. In Subtest C the subject is presented a short rhythm pattern from the tape recorder and

asked to accurately reproduce it by hitting together two square wooden blocks. The final subtest (Subtest D) involves the subject comparing two rhythms and determining if they are alike. The last two subtests involve rhythms ranging from three to seven notes.

The rationale for using these four subtests as a measure of rhythmical ability is that most Western music is constituted of rhythm involving a more or less steady pulse (or beats) on which are superimposed rhythmic patterns. It is necessary for a person with some rhythmic ability to be able, to some degree, to hear the difference between like and unlike tempi and rhythms in order to understand or enjoy the rhythmic qualities of music. Also, for a person to produce music, that person must be able to produce, with some degree of accuracy, rhythmic patterns and recurring pulses. This is especially important when one is engaged in music-making with others.

It is quite possible that a subject might have varying degrees of ability to either accurately hear or reproduce these two rhythmic elements. The test was designed to take into account these factors. The two subtests (A and D) involving a comparison of two stimuli require a verbal response from the subject. The other subtests (B and C) require a nonverbal one.

Research into Test Reliability and Validity

The original form of the test was given to 16 four-year-old subjects. The split-half reliability was .91 when the Spearman-Brown prophecy formula was employed to approximate the reliability of the full length of the test.

In an effort to make improvements in the test, the writer modified nine of the test items which had unacceptable levels of discrimination or item difficulty. The new form of the test was administered to 10 four-year-old subjects to determine the item difficulty and discrimination of the modified items. The new test form was also administered to 36 four-year-old subjects who participated in an experiment to determine the effects of training on posttest scores. The split-half reliabilities of the composite test scores ranged from .86 to .89 for the four-year-olds who took the new test. The split-half reliability for scores of the experimental and control group subjects in the experiment was the same (.86) for both pre- and posttest when the scores of the subjects were taken as a single group; however, there was an intervening variable between the two tests for two-thirds of the subjects.

The test-retest reliability for the 36 subjects in the experiment was .45 for all subjects taken as a single group. This correlation itself was not significant at or beyond the .05 level. This may be accounted for in part by the intervening variables. The test-retest reliability for the one-third of the subjects who did not have exposure to the intervening variables was .75. This correlation was significant. This is not as high a correlation as the split-half reliability of the scores of the subjects in the experiment, but it is a fairly high correlation when the consideration is made that this datum (.75) was derived from the performances of only 12 subjects.

The validity of both forms of the test was investigated. The scores of 32 fourth-graders who took the old form of the researcher's

test were correlated with their scores on the standardized test the Musical Aptitude Profile (MAP) of Edwin Gordon. The composite scores from the researcher's test correlated highest with the following scores which were derived from scores yielded from an administration of the MAP: composite for "Rhythmic Imagery" (.65), "Rhythmic Imagery II"--meter (.61), "Rhythmic Imagery I"--tempo (.60), and composite for the whole MAP (.51). Each of these correlations was significant at or beyond .001. The correlations of scores between the researcher's test and the composite MAP scores for "Musical Sensitivity" and "Tonal Imagery" were .45 and .25 respectively, but only the first correlation was significant at the .05 level.

The validity of the new form of the test was investigated through the correlation of scores received by 10 kindergarten-age subjects on the test with rankings of their "musical rhythmical ability" provided by their classroom and music teacher. The ranking of the music teacher correlated with composite test scores at .78, and the correlation of the classroom teacher correlated at .66. The Spearman rank-order formula was used in the computation.

The Effects of Training Experiment

The 36 subjects used in the experiment were randomly selected from a pool of 62 four-year-old preschool students at five preschool centers in Montgomery County, Maryland. The subjects were separated into three subgroups which contained an equal number of boys and girls. Each subgroup also had an equal distribution of subjects representing young, middle-aged, and older four-year-olds.

One subgroup was randomly selected to act as a control group which had no contact with the researcher between testings (Group C). Another group was randomly selected to have contact with the researcher (Group CC). The researcher met with these subjects for 20 sessions between the pre- and posttest. Each session lasted about 12 minutes and involved nonmusical activities led by the researcher. The reason for having this group was to determine whether the researcher's personal contact would affect performance on the posttest. The experimental group (Group X) received 20 training sessions which lasted about 12 minutes each. The sessions for Group CC and Group X were conducted within the same period of time which was one month. The training for the experimental group consisted of equal participation by these subjects in the following:

1. Determining whether two tempi produced by a metronome were the same or different
2. Reproducing on a percussion instrument various tempi produced by a metronome
3. Duplicating rhythm patterns produced by the researcher
4. Determining whether two rhythm patterns produced by the researcher were the same or different

These activities represent the four abilities measured by the rhythm test. None of the training stimuli were exactly like the test stimuli. Four training sessions involved less structured activity like marching.

Hypothesis 1, that training would improve scores, was tested using analysis of covariance. An F ratio was computed for each subtest and for the composite test scores. None of the computed ratios was sufficient to reject the null hypothesis at or beyond the .05 level of

confidence.

Hypothesis 2 stated that the test scores of control subjects having contact with the researcher would show a greater, though insignificant, increase from pre- to posttest than the scores of the control subjects who did not have contact with the researcher. The fact that the subjects who did not have contact with the researcher between testings had a greater mean increase was not consistent with this hypothesis.

Hypothesis 3, that neither boys nor girls in the four-year-old age-group would perform significantly better on the rhythm test, was a valid hypothesis since the F ratio generated through an analysis of variance was not significant at the .05 level.

Performances on the Various Subtest Items

The pretest data of the 36 subjects who participated in the experiment were analyzed to discover information about the performances of four-year-olds on the various rhythmical stimuli in the four subtests. Pretest data was used because there were intervening variables before the posttest.

In Subtest A, in which a subject determines if two tempi are the same, the subjects did significantly better on the items which had different rates when the slower rate came first. The chi-square value was significant at .005. Also, there was a -.42 correlation between the speed of stimuli in items of Subtest A having identical tempi and the number of correct responses. This indicates, though only to a small degree, that the faster the stimuli of items having identical tempi, the

more subjects tend to believe they are hearing different stimuli. All of the correlations in the present discussion were themselves significant at the .05 level.

Subtest B requires the subject to hear a given tempo and then reproduce it accurately. In this subtest the correlation was computed between the number of correct responses to the various items and the tempo of the items. The correlation was .77. This indicates that the faster the tempo between 60 to 132 beats per minute, the easier it is for a subject to reproduce that stimulus accurately.

Subtest C involves the subject hearing and then reproducing a rhythm. The correlation between the number of taps in an item and the mean number of correct tap-responses of the subjects to those items was found to be -.94. This strongly suggests that a four-year-old's success in getting the correct number of taps to an item increases as the number of taps decrease in rhythms having from three to seven taps per item.

Other Findings

The following are findings from various other phases of the investigation:

1. All three groups involved in the experiment had higher mean scores on the posttest.

2. A *t*-test was applied to the pre- and posttest scores of each group in the experiment taken separately. The improvement in the scores of Group X was significant beyond .01, and Group CC had a significant improvement which was beyond .05. Group C improved, but the improvement was at the .10 level and not significant.

3. The scores of each separate age-group's data were subjected to an analysis of covariance to determine if there was a

significant difference in the performances among the three age-subgroups on the posttest. Analyses were made for the composite scores and scores for each of the subtests. None of the F values derived from the analyses revealed a significant difference in performance.

4. The correlation of age in months with pretest scores was .46 for the composite test scores of the four-year-olds in the experiment. This indicates that there is a small relationship between age and performance on the rhythm test among the four-year-olds. This finding is consistent with the findings of other researchers who have studied the rhythmic responses of preschool-age subjects.

5. The data of the scores of the subjects in the experiment were compared with the data of the kindergarten subjects who also took the new version of the test. An analysis of variance revealed no significant difference in performance between the two age-levels.

6. The data of the scores of the 32 fourth-grade students and 16 four-year-olds who took the old form of the test were compared. An analysis of variance revealed that the subjects in the fourth-grade group performed significantly better on the test. The level of significance was beyond .001.

Conclusions

The data from this research indicate that the rhythmic ability of four-year-olds is not subject to improvement after one month of training. If training will significantly improve the rhythmic ability of four-year-olds at all, then, the results of the study indicate that the improvement will be a slower and more gradual process than can be accomplished in daily training for one month.

A longer period of training might show completely different results. Also, the type of training and the way it is implemented may show different results. In the opinion of the researcher, the training sessions used in this study were rather uninteresting, at times, for some of the experimental subjects.

The personal contact that the researcher had with the subjects between testings did not seem to be an important factor in the performance on the posttest. This indicates that a researcher can conduct data gathering activities similar to those of this study with a minimum of contact with the subjects before the data gathering activities without jeopardizing the outcome of the experiment. The researcher found that only a few of the subjects he encountered were shy, afraid, or extremely withdrawn in the presence of the experimenter when they had their first encounter with him. Some teachers commented that they were quite surprised that the children participated in the experiment so willingly. The giving of candy, the researcher subjectively feels, was responsible for this.

Apparently, there appears to be no significant difference in the rhythmic ability of boys or girls at age four. This confirms the findings of other researchers.

The rhythm test as a whole appears to have acceptable levels of test reliability, and the results from the study of the test's validity seem promising. There are weaknesses in the battery. For example, some of the subtests have low reliabilities. The test does seem to have potential for being developed into a standardized battery which can be used for more definitive research findings. The test does produce fairly reliable results with young children when all the factors and problems related to testing this age-group are considered.

Need for Further Research

The test itself needs further refinement and study. Also, it needs to be administered to larger groups of children. Subtest A, for

example, might be modified so that some of the items having stimuli of different rates are made even more divergent. This may raise the reliability of this subtest.

The reliability and validity data on Subtest *D*, especially the data from the performance by the kindergarten subjects, indicate that this subtest needs further modification and refinement. Findings seem to indicate that increasing the length (number of notes) of the items may produce better reliability of performance since the length of the items correlated at .66 with the number of correct responses to the items; however, the correlation was not significant.

The more reliable responses of the subjects to stimuli requiring nonverbal responses indicate that a four-year-old may be more attentive to stimuli when he knows he can respond to them nonverbally. This point deserves further research.

This study indicates that rhythmic ability is possibly on some sort of continuum between preschool children and fourth-grade or older children. Further research is needed to determine to what degree and how rhythmic ability is a function of age. A longitudinal study with the test would seem quite appropriate.

Finally, the results of the study indicate that testing the responses of four-year-old children to rhythmic stimuli in a controlled situation is plausible using more or less conventional techniques. Further research with four-year-olds and, perhaps, even younger children would seem fruitful using techniques developed in this study.

An enormous amount of research is needed in all phases of the rhythmic element of music and the responses to it by young children.

BIBLIOGRAPHY

Books

Akers, Milton E. "Prologue: The Why of Early Childhood Education." Early Childhood Education. Seventy-fifth Yearbook of the National Society for the Study of Education, Part II. Chicago: The University of Chicago Press, 1972.

Bloom, Benjamin S.; Davis, Allison; and Hess, Robert. Compensatory Education for Cultural Deprivation. New York: Holt, Rinehart and Winston, Inc., 1960.

Bruning, James L., and Kintz, B. L. Computational Handbook of Statistics. Glenview, Illinois: Scott, Foresman and Company, 1968.

Christianson, Helen. Bodily Rhythmic Movements of Young Children in Relation to Rhythm in Music. No. 736: Contributions to Education. New York: Bureau of Publications, Teachers College, Columbia University, 1938.

Evans, Ellis D. Contemporary Influences in Early Childhood Education. New York: Holt, Rinehart and Winston, Inc., 1971.

Farnsworth, Paul R. The Social Psychology of Music. Ames, Iowa: The Iowa State University Press, 1969.

Glenn, Neal E., and Turrentine, Edgar M. Introduction to Advanced Study in Music Education, Dubuque, Iowa: Wm. C. Brown Publishers, 1968.

Gordon, Edwin. The Psychology of Music Teaching. Englewood Cliffs, N. J.: Prentice-Hall, Inc., 1971.

_____. A Three-Year Longitudinal Predictive Study of the Musical Aptitude Profile. Vol. V: Studies in the Psychology of Music. Iowa City, Iowa: University of Iowa, 1968.

Hattwick, Melvin S. "A Genetic Study of Differential Pitch Sensitivity," in The Measurement of Musical Development II. Vol. XI, No. 2: University of Iowa Studies: Studies in Child Welfare. Iowa City, Iowa: University of Iowa, 1935.

Hicks, James A. The Acquisition of Motor Skill in Young Children. Vol. IV, No. 5: University of Iowa Studies: Studies in Child Welfare. Iowa City, Iowa: University of Iowa, 1931.

Jersild, Arthur T., "Music." Child Development and the Curriculum. Thirty-eighth Yearbook of the National Society for the Study of Education, Part I. Bloomington, Illinois: Public School Publishing Company, 1939.

_____, and Bienstock, Sylvia F. Development of Rhythm in Young Children. Child Development Monographs, XXII, 1935.

LaCrosse, Robert E., ed. Early Childhood Education Directory. New York: R. R. Bowker Company, 1971.

Lehman, Paul R. Tests and Measurements in Music. Englewood Cliffs, N. J.: Prentice-Hall, Inc., 1968.

Lundin, Robert W. An Objective Psychology of Music. New York: The Ronald Press Company, 1967.

Seashore, Carl E. Psychology of Music. New York: McGraw-Hill Book Company, Inc., 1938.

Sigel, Irving E. "Developmental Theory and Preschool Education: Issues, Problems and Implications." Early Childhood Education. Seventy-fifth Yearbook of the National Society for the Study of Education, Part II. Chicago: The University of Chicago Press, 1972.

Updegraff, Ruth; Heiliger, Louise; and Learned, Janet. "The Effect of Training Upon the Singing Ability and Musical Interest of Three-, Four-, and Five-year-old Children," in Studies in Preschool Education I. Vol. XIV, No. 1: University of Iowa Studies: Studies in Child Welfare. Iowa City, Iowa: University of Iowa, 1938.

Wellman, Beth. The Development of Motor Co-ordination in Young Children. Vol. III, No. 4; University of Iowa Studies: Studies in Child Welfare. Iowa City, Iowa: University of Iowa, 1926.

Whybrew, William E. Measurement and Evaluation in Music. 2nd ed. Dubuque, Iowa: Wm. C. Brown Company Publishers, 1971.

Williams, Harold M. "Studies in the Measurement of Musical Development," in The Measurement of Musical Development. Vol. VII, No. 1: University of Iowa Studies: Studies in Child Welfare. Iowa City, Iowa: University of Iowa, 1932.

Periodicals

Colby, Martha G. "Instrumental Reproduction of Melody by Pre-School Children." Journal of Genetic Psychology, XLVII (December 1935), 413-430.

Drexler, Edith N. "A Study of the Development of the Ability to Carry a Melody at the Pre-school Level." Child Development, IX (September, 1938), 319-332.

Fowler, William. "Cognitive Learning in Infancy and Early Childhood." Psychological Bulletin, LIX (March, 1962), 135-152.

Friend, Ruby S. "Influences of Heredity and Musical Environment on the Scores of Kindergarten Children on the Seashore Measures of Musical Ability." Journal of Applied Psychology, XXIII (June, 1939), 347-357.

Fullard, William G., Jr. "Operant Training of Aural Musical Discriminations with Preschool Children." Journal of Research in Music Education, XV (Fall, 1967), 201-209.

Harrington, Charles J. "An Investigation of the Experimental Version Primary Level Musical Aptitude Profile for Use with Second and Third Grade Students." Journal of Research in Music Education, (Winter, 1969), 193-201.

Hattwick, Melvin S. "The Role of Pitch and Pitch Range in the Singing of Preschool, First Grade, and Second Grade Children." Child Development, IV (December, 1933), 281-291.

Heinlein, Christian Paul. "A New Method of Studying the Rhythmic Responses of Children Together with an Evaluation of the Method of Simple Observation." Journal of Genetic Psychology, XXXVI (June, 1929), 205-227.

Hissem, Irene. "A New Approach to Music for Young Children." Child Development, IV (December, 1933), 308-317.

Jersild, Arthur T., and Bienstock, Sylvia F. "The Influence of Training on the Vocal Ability of Three-Year-Old Children." Child Development, II (December, 1931), 272-291.

_____, and Bienstock, Sylvia F. "A Study of the Development of Children's Ability to Sing." Journal of Educational Psychology, XXV (October, 1934), 481-503.

McGinnis, Esther. "Seashore's Measures of Musical Ability Applied to Children of the Pre-School Age." American Journal of Psychology, XL (October, 1928), 620-623.

Seltzer, Serkphine. "A Measure of the Singing and Rhythmic Development of Preschool Children." Journal of Educational Psychology, XXVII (September, 1936), 417-424.

Tarrell, Vernon. "An Investigation of the Validity of the Musical Aptitude Profile." Journal of Research in Music Education, XIII

(Winter, 1965), 195-206.

Vance, Thomas F., and Grandprey, Medora B. "Objective Methods of Ranking Nursery School Children on Certain Aspects of Musical Capacity." Journal of Educational Psychology, XXII (November, 1931), 577-585.

Wight, Minnie G. "The Effect of Training on Rhythmic Ability and Other Problems Related to Rhythm." Child Development, VIII (June, 1937), 159-172.

Theses and Dissertations

Broadhead, Mary Caryl. "Musical Opportunities in the Home and Their Relationship to the Musical Achievement of Kindergarten Children." Unpublished Master's thesis, Cornell University, 1958.

DeYarman, Robert M. "Experimental Analysis of the Development of Rhythmic and Tonal Capabilities of Kindergarten and First Grade Children." Unpublished Ph.D. dissertation, University of Iowa, 1971.

Fosha, Leon. "A Study of the Validity of the Musical Aptitude Profile." Unpublished Ph.D. dissertation, University of Iowa, 1960.

Kirkpatrick, William C., Jr. "Relationships Between the Singing Ability of PreKindergarten Children and Their Home Musical Environment." Unpublished Ph.D. Dissertation, University of Southern California, 1961.

Reynolds, George E. "Environmental Sources of Musical Awakening in Pre-School Children." Unpublished Ed.D. Dissertation, University of Illinois, 1960.

Shelton, John Stanley. "The Influence of Home Musical Environment Upon Musical Response of First-Grade Children." Unpublished Ed.D. dissertation, George Peabody College for Teachers, 1965.

Shull, Dorothy. "A Study of the Influence of the Musical Environment in the Home on the Musicality of Selected Kindergarten Children." Unpublished Master's thesis, Northwestern University, 1953.

Smith, Robert Barton. "A Study of the Effect of Large-Group Vocal Training on the Singing Ability of Nursery School Children." Unpublished Ed.D. Dissertation, University of Illinois, 1960.

Tests

Bentley, Arnold. Measures of Musical Ability. New York: October House, Inc., 1966.

Gordon, Edwin. Musical Aptitude Profile Manual. Boston: Houghton Mifflin Company, 1965.

APPENDIX A

TEST ACCLIMATIZATIONS

Test A

1. The experimenter greets the subjects.
2. The subjects take their seats.
3. "I have some candy for you because you are so nice."
4. The experimenter gives each child one piece of M&M candy.
5. The experimenter holds up a metronome.
6. "This is a music clock. It ticks like a clock."
7. The experimenter lets the clock tick eight times at 63 beats per minute (bpm).
8. "This is a slow clock."
9. The experimenter lets the clock tick eight more times at 63 bpm.
10. The next one is a fast sounding clock."
11. Eight ticks are played at 184 bpm.
12. "Listen carefully and see if these two clocks sound the same to you."
13. Eight ticks are played at 100 bpm.
14. "Here's the other."
15. Eight ticks are played at 100 bpm.
16. "They sound the same, don't they? Here is the sound of two more clocks. Listen carefully and see if they sound the same or different."
17. Eight ticks are played at 63 bpm.
18. "Here's the other."

19. Eight ticks are played at 184 bpm.

20. "These two clocks don't sound the same, do they?"

21. "This is a music box."

22. The experimenter points to the tape recorder.

23. "It goes tick-tock like a clock."

24. The experimenter plays Examples A and B from Test A.

25. "Now I am going to play for you the sound of two clocks. Listen carefully and see if they sound the same or if they sound different."

26. The experimenter plays Example A from Test A.

27. "They sound the same, don't they?"

28. "Now listen to these two clocks and see if they are the same or different."

29. The experimenter plays Example B from Test A.

30. "They sound different, don't they?"

31. At this point each child is permitted to answer "same" or "different" after hearing both examples. The presentations of the examples is randomized. No child responds to both examples concurrently. Correct responses are rewarded with the word "good." The tester says "no" when a response is incorrect.

32. "We have heard some clocks today. Some sound the same and some sound different. You have been so much fun. I am going to give each of you some candy because you are all good and nice."

33. The experimenter gives each child two pieces of M&M candy and sends the children back to their class.

Test B

1. The experimenter greets the subjects.

2. The subjects take their seats.

3. "I have some candy for you because you are so nice."

4. The experimenter gives each child one piece of M&M candy.

5. The experimenter holds up a metronome.

6. "This is a music clock. It ticks like a clock."

7. The experimenter lets the metronome tick eight times at 56 bpm.

8. "Some clocks go slow like this."

9. Eight ticks are played at 56 bpm.

10. "Some clocks go fast like this."

11. Eight ticks are played at 152 bpm.

12. "Now I am going to show you some music blocks."

13. The experimenter holds up the wooden blocks.

14. "We can hit the music blocks together and make them sound like the tick-tock of a clock."

15. The experimenter hits the blocks together at about 80 bpm.

16. "Let's see if you can make the music blocks sound like a clock."

17. Each child is given the blocks and allowed to experiment. The experimenter stops each child after about eight beats or so.

18. "Now let's see if we can stay together with the music clock. We will hit the blocks at the same time the clock ticks."

19. The experimenter demonstrates by hitting the blocks together in tempo with the metronome which is set at 83 bpm.

20. "Let's each of us try it."

21. The experimenter lets each child hit the blocks with the metronome set at 83 bpm. Each child is stopped after about ten ticks.

22. "This is a music box."

23. The experimenter points to the tape recorder.

24. "It goes tick-tock like a clock. Listen to it."

25. The experimenter plays the first eight metronome ticks of Example A in Test B.

26. "I am going to hit the blocks along with the clock. I am going to keep hitting until I hear the bell ring."

27. The experimenter plays Example A and taps along with the metronome. The tapping is not stopped until the bell rings.

28. "What did I do when I heard the bell? I stopped."

29. "Now you can try it. Keep hitting until you hear the bell. Try to hit like the clock is ticking."

30. Each child has the opportunity to do both examples; however, no child performs both examples concurrently. Order is random.

31. "We have heard some clocks today. You have hit the blocks yourself along with the clocks. You have been so much fun. I am going to give each of you some candy because you are all good and nice."

32. The experimenter gives each child two pieces of M&M candy and sends the children back to their class.

Test C

1. The experimenter greets the subjects.

2. The subjects take their seats.

3. "I have some candy for you because you are so nice."

4. The experimenter gives each child one piece of M&M candy.

5. "Now I am going to show you some blocks."

6. The experimenter holds up the wooden blocks.

7. "We can hit the blocks together and make a sound."

8. The experimenter makes several sounds by hitting the blocks.

9. "Let's each of us make a sound."

10. Each child is given the blocks and allowed to hit the blocks several times.

11. "Now I am going to ask you to do something. Listen and see if you can do it."

12. The experimenter hits the blocks together twice at about 80 bpm.

13. "Now you can do it."

14. Each child gets the opportunity to respond to the experimenter's

stimulus. If the child has the correct number of taps and about the same tempo, then, the experimenter says "good." If either the number of taps or the tempo is incorrect the child is given another opportunity to match the stimulus.

15. "This is a music box."

16. The experimenter points to the tape recorder.

17. "It can make sounds. We just made some sounds with the blocks. Listen to the music box."

18. The experimenter plays Examples A and B of Test C.

19. "Now I am going to listen to the music box. After I hear the music box I am going to hit the blocks like the music box sounds."

20. The experimenter plays Example A from Test C and then hits the blocks together with the correct tempo and rhythm of the stimulus.

21. "Now you can try it. Listen carefully to the music box and then hit the blocks like what you heard."

22. Each child is given the opportunity to hit the blocks together after hearing the stimulus of Example A.

23. "That was fun. Now let's hear the music box play again. After it plays, each of you can make the same sound on your blocks."

24. Each child is given the opportunity to hit the blocks together after hearing Example B; however, this time correct responses are rewarded with the word "good." Subjects making incorrect responses are told so, and they are given another opportunity to respond.

25. "We have heard some sounds made by the music box today. We hit our blocks together like the music box sounded. You hit the blocks after you heard the music box. You have been so much fun. I am going to give each of you some candy because you are all good and nice."

26. The experimenter gives each child two pieces of M&M candy and sends the children back to their class.

Test D

1. The experimenter greets the subjects.

2. The subjects take their seats.

3. "I have some candy for you because you are so nice."

4. The experimenter gives each child one piece of M&M candy.

5. "Now I am going to show you some music blocks."

6. The experimenter holds up the wooden blocks.

7. "We can hit the music blocks together and make a sound."

8. The experimenter hits the blocks together several times.

9. "Let's each of us make a sound."

10. Each child is given the blocks and allowed to experiment.

11. "Now I am going to ask you to do something. Listen to these two sounds."

12. The experimenter plays four taps at about 83 bpm. After a pause of about two seconds, he plays only two taps at about the same rate.

13. "Were these two sounds the same or different?"

14. The experimenter gives the subjects time to respond. Correct responses are rewarded with "good."

15. "Now listen to these two music sounds and see if they are the same or different."

16. The experimenter plays the stimulus of two taps at about 83 bpm. After about two seconds, the stimulus is repeated.

17. "Were these two sounds the same or different?

18. Correct responses are rewarded with "good."

19. "We know that some music sounds are the same and some are different. We heard two sounds that were the same and two that were different."

20. "This is a music box."

21. The experimenter points to the tape recorder.

22. "It makes music sounds like we just made. Listen to some of the sounds."

23. The experimenter plays Examples A and B from Test D.

24. "Now I am going to play some sounds from the music box. This time listen carefully and see if they are the same sounds or different sounds."

25. The experimenter plays Example A.

26. "These two sounds are the same."

27. "Now listen to this one."

28. The experimenter plays Example B.

29. "These two sounds are different."

30. "Now you try it."

31. At this point each child is permitted to answer "same" or "different" after hearing Examples A and B. The presentation of the stimuli is random. No child responds to both examples concurrently. Correct responses are rewarded with the word "good." The tester says "no" when a response is incorrect.

32. "We have heard some sounds today. Some sounds are the same, but some sounds are different. You told me if they sounded the same to you, and you told me if they sounded different to you. You have been so much fun. I am going to give each of you some candy because you are all good and nice."

33. The experimenter gives each child two pieces of M&M candy and sends the children back to their class.

APPENDIX B

THE RHYTHM TEST

Test A: Like-Unlike Tempi

Preliminaries

1. The experimenter greets the subject.

2. The subject takes his seat.

3. "I have some candy for you because you are so nice."

4. The experimenter gives the subject one piece of M&M candy.

5. "We heard some clocks this morning. Do you remember? Some of the clocks sounded the same and some sounded different. Now I am going to play the sound of two clocks. Listen carefully and see if they sound the same or if they sound different."

6. The experimenter plays Example A from Test A.

7. If the subject makes a correct response, the experimenter says the following:

   "That was right. I am going to let you hear these same two clocks again. They are the same."

   If the subject makes an incorrect response the experimenter says:

   "No, that was not right. I am going to let you hear these same two clocks again. They are the same."

8. The experimenter plays Example A again.

9. "I am going to play the sound of two more clocks. Listen carefully and see if they sound the same or if they sound different."

10. The experimenter plays Example B.

11. If the subject makes a correct response, the experimenter says the following:

"That was right. I am going to let you hear these same two clocks again. They are different."

If the subject makes an incorrect response the experimenter says:

"No, that was not right. I am going to let you hear these same two clocks again. They are different."

12. The experimenter plays Example B.

13. "Now listen carefully. I am going to play some clocks, and you tell me if they sound the same or different."

14. The researcher plays the ten test items and records the subject's responses.

15. The experimenter talks for about one-half minute with the subject after the test is completed. The subject is then given two pieces of M&M candy and sent back to his group.

Test A

The following is the practice items and test items as they were recorded on the tape. The abbreviation "bpm" stands for beats per minute. The items in parentheses are the modified forms of the items.

A. 100 bpm followed by 100 bpm

B. 63 bpm followed by 184 bpm

1. 66 bpm followed by 42 bpm

2. 72 bpm followed by 96 bpm

3. 60 bpm followed by 60 bpm (138 bpm followed by 138 bpm)

4. 96 bpm followed by 120 bpm

5. 120 bpm followed by 96 bpm (112 bpm followed by 88 bpm)

6. 132 bpm followed by 132 bpm

7. 88 bpm followed by 88 bpm

8. 112 bpm followed by 88 bpm (120 bpm followed by 96 bpm)

9. 80 bpm followed by 80 bpm

10. 104 bpm followed by 104 bpm

Test B: Continue Taps

Preliminaries

1. The experimenter greets the subject.
2. The subject takes his chair.
3. "I have some candy for you because you are so nice."
4. The experimenter gives the subject one piece of M&M candy.
5. "We heard some clocks this morning. Do you remember? You hit the blocks yourself along with the clocks. I am going to hit the blocks together with the clock. I am going to keep hitting until I hear the bell."
6. The experimenter plays Example A from Test B and hits the blocks together.
7. "Now you can hit the blocks together with the clock."
8. The experimenter plays Example A and lets the child hit the blocks together.
9. The experimenter plays Example B and lets the child hit the blocks together.
10. "Now listen carefully. I am going to play some more clocks. You keep hitting until you hear the bell.
11. The researcher plays the ten test items and tape records the subject's responses.
12. The experimenter talks for about one-half minute with the subject after the test is completed. The subject is then given two pieces of M&M candy and sent back to his group.

Test B

The following is the practice items and test items as they were recorded on the tape. The abbreviation "bpm" stands for beats per minute.

A. 93 bpm

B. 93 bpm

1. 72 bpm

2. 80 bpm

3. 66 bpm

4. 120 bpm

5. 96 bpm

6. 60 bpm

7. 132 bpm

8. 88 bpm

9. 112 bpm

10. 104 bpm

Test C: Rhythm Duplication

Preliminaries

1. The experimenter greets the subject.

2. The subject takes his seat.

3. "I have some candy for you because you are so nice."

4. The experimenter gives the subject one piece of M&M candy.

5. "We heard some sounds made by the music box this morning. Do you remember? You hit the blocks yourself like the music box sounded. I am going to listen to the music box. After I hear the music box, I am going to hit the blocks like the music box sounds."

6. The experimenter plays Example A from Test C and hits the blocks together after the stimulus.

7. "Now you can hit the blocks together just like the music box. Wait until after the music box, then hit."

8. The experimenter plays Example A and lets the child hit the blocks.

9. The experimenter plays Example B and lets the child hit the blocks.

10. "Now listen carefully. I am going to play some sounds from the music box. You hit the blocks together just like the music box sounds. Wait for the music box to go first."

11. The researcher plays the ten test items and tape records the subject's responses.

12. The experimenter talks for about one-half minute with the subject after the test is completed. The subject is then given two pieces of M&M candy and sent back to his group.

Test C

The following is the practice items and test items as they were recorded on the tape. Modified items are in parentheses. A modified form of musical notation will be used to present the test items. The following figures will be used: o/ (half note), x/ (quarter note), x/⌐ (eighth note), x/= (sixteenth note).

A. M.M. x/=93 (played on claves)

$\frac{4}{4}$ x/⌐x/⌐ x/ x/ x/

B. M.M. x/=93 (played on rhythm sticks)

$\frac{4}{4}$ x/⌐x/⌐ x/ x/ x/

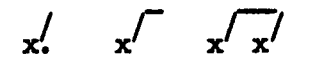
1. M.M. x/=88 (played on tom-tom)

$\frac{5}{4}$ x/ x/ x/⌐x/⌐ x/ x/ $\frac{5}{4}$ (x/ x/ x/⌐x/⌐ o/)

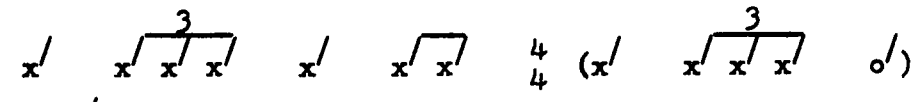
2. M.M. o/=80 (played on marimba)

$\frac{3}{2}$ x/ x/ x/. x/⌐ o/

3. M.M. x/=66 (played on piano--Middle-C)

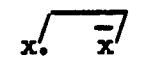


4. M.M. $x^{/}$=60 (played on rhythm sticks)

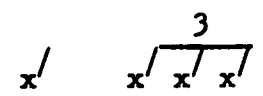


5. M.M. $x^{/}$=72 (played on snare drum)

6. M.M. $x^{/}$=112 (played on tambourine)



7. M.M. $x^{/}$=96 (played on electric piano--E above Middle-C)



8. M.M. $x^{/}$=120 (played on xylophone--G above Middle-C)

9. M.M. $x^{/}$=132 (played on cowbell)

10. M.M. $x^{/}$=104 (played on claves)

Test D: Like-Unlike Rhythms

Preliminary

1. The experimenter greets the subject.
2. The subject takes his seat.

3. "I have some candy for you because you are so nice."

4. The experimenter gives the subject one piece of M&M candy.

5. "We heard some sounds made by the music box this morning. Do you remember? Some of the sounds were the same and some were different. Now I am going to play two sounds from the music box. Listen carefully and tell me if they sound the same or different."

6. The experimenter plays Example A from Test D.

7. If the subject makes a correct response, the experimenter says the following:

   "That was right. I am going to let you hear these sounds again. They are the same."

   If the subject makes an incorrect response the experimenter says:

   "No, that was not right. I am going to let you hear these sounds again. They are the same."

8. The experimenter plays Example A again.

9. "I am going to play two more sounds. Listen carefully and tell me if they sound the same or different."

10. The experimenter plays Example B.

11. If the subject makes a correct response, the experimenter says the following:

    "That was right. I am going to let you hear these sounds again. They are different."

    If the subject makes an incorrect response the experimenter says:

    "No, that was not right. I am going to let you hear these sounds again. They are different."

12. The experimenter plays Example B again.

13. "Now listen carefully. I am going to play some sounds from the music box. Tell me if they are the same or different."

14. The researcher plays the ten test items and records the subject's responses.

15. The experimenter talks for about one-half minute with the subject after the test is completed. The subject is then given two pieces

of M&M candy and sent back to his group.

Test D

The following is the practice items and test items as they were recorded on the tape. Modified items are presented below the original items and enclosed within parentheses. A modified form of musical notation will be used to represent the test items. The following figures will be used: o/ (half note), x/ (quarter note), x/⁀ (eighth note), x/= (sixteenth note). Note that item "2" was rerecorded in the original notation but using claves instead of the marimba.

A. M.M. x/=93 (played on tom-tom)

$\frac{3}{4}$ x/ $\overset{3}{x/\,x/\,x/}$ x/ $\frac{3}{4}$ REPEATED UNCHANGED

B. M.M. x/=93 (played on tom-tom)

$\frac{3}{4}$ x/⁀x/ x/⁀x/ x/ $\frac{3}{4}$ x/ $\overset{3}{x/\,x/\,x/}$ x/

1. M.M. x/=104 (played on tambourine)

$\frac{5}{4}$ x/ x/⁀x/ x/ x/ x/ $\frac{5}{4}$ REPEATED UNCHANGED

2. M.M. x/=72 (played on marimba--Middle-C)

$\frac{3}{4}$ x/⁀ x/ x/ x/⁀ $\frac{3}{4}$ x/ x/ x/⁀x/

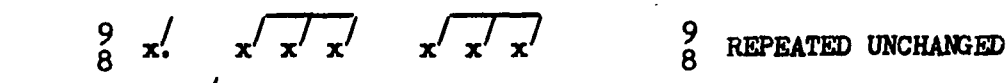
(The modified form was rerecorded as is using claves)

3. M.M. x/=60 (played on tom-tom)

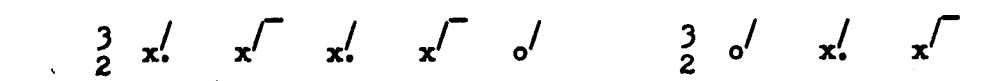
$\frac{2}{4}$ $\overset{3}{x/\,x/\,x/}$ x/ $\frac{2}{4}$ REPEATED UNCHANGED

$\frac{2}{4}$ ($\overset{3}{x/\,x/\,x/}$ x/=x/=) $\frac{2}{4}$ (REPEATED UNCHANGED)

4. M.M. $x^{\prime}$=80 (played on snare drum)

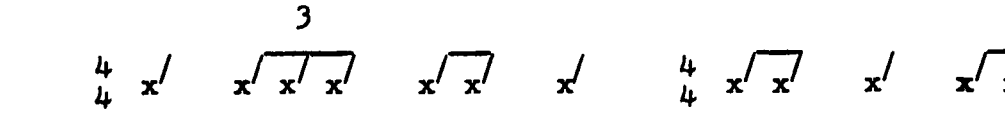


5. M.M. $o^{\prime}$=96 (played on xylophone--G above Middle-C)

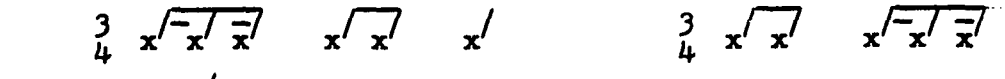


6. M.M. $x^{\prime}$=66 (played on electric piano--E above Middle-C)

7. M.M. $x^{\prime}$=120 (played on claves)

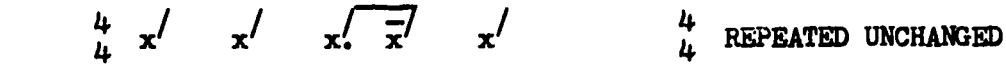


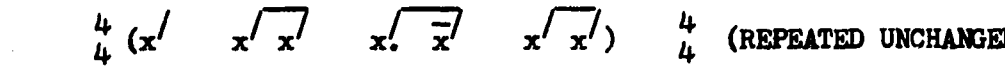
8. M.M. $x^{\prime}$=132 (played on rhythm sticks)



9. M.M. $x^{\prime}$=88 (played on cowbell)

10. M.M. $x^{\prime}$=112 (played on piano--A below Middle-C)





APPENDIX C

SCHEDULE OF ACTIVITIES FOR GROUP CC DURING THEIR EXPOSURE TO THE RESEARCHER BETWEEN PRETEST AND POSTTEST

First Week

Monday-----Researcher read the child's book *Old Hat New Hat*.
Tuesday----Researcher showed Canadian and American coins and discussed money.
Wednesday--Researcher read the child's book *The Busy Bulldozer*.
Thursday---Researcher read the child's book *A Garden is Good*.
Friday-----Researcher read the child's book *The Ear Book*.

Second Week

Monday-----Researcher read the child's book *Hand, Hand, Fingers, Thumb*.
Tuesday----Researcher showed and discussed playing cards.
Wednesday--Researcher read the child's book *Go Away, Dog*.
Thursday---Researcher read the child's book *Bears on Wheels*.
Friday-----Researcher read the child's book *Mother Goose*.

Third Week

Monday-----Researcher showed and discussed a road map.
Tuesday----Researcher read the child's book *Birds*.
Wednesday--Researcher read the child's book *Mr. Brown Can Moo! Can you?*
Thursday---Researcher read the child's book *Pony Twins*.
Friday-----Researcher read the child's book *The Gingerbread Man*.

Fourth Week

Monday-----Researcher read the child's book *Toy Train*.
Tuesday----Researcher discussed numbers and counting.
Wednesday--Researcher read the child's book *Three Billy Goats Gruff*.
Thursday---Researcher read the child's book *Fly High*.
Friday-----Researcher read the child's book *Our Animal Friends*.

APPENDIX D

SCHEDULE AND EXPLANATION OF TRAINING ACTIVITIES FOR GROUP X BETWEEN PRETEST AND POSTTEST

Training Schedule

First Week

Monday-----Training B, Session I
Tuesday----Training A, Session I
Wednesday--Training D, Session I
Thursday---Training C, Session I
Friday-----Marching to recordings of band marches

Second Week

Monday-----Training B, Session II
Tuesday----Training C, Session II
Wednesday--Clapping to recordings of band marches
Thursday---Training A, Session II
Friday-----Training D, Session II

Third Week

Monday-----Training B, Session III
Tuesday----Walking to the rhythm of simple children's songs played on a phonograph
Wednesday--Training C, Session III
Thursday---Training D, Session III
Friday-----Training A, Session III

Fourth Week

Monday-----Training C, Session IV
Tuesday----Clapping the rhythm of simple children's songs played on a phonograph
Wednesday--Training A, Session IV
Thursday---Training D, Session IV
Friday-----Training B, Session IV

The following is a discussion of the development and use of each of the four types of training. Included also are the activities for each of the sessions.

TRAINING A

Each of the four training sessions for Test A consist of the presentation of eight pairs of tempi to the experimental subjects. Four pairs have different tempi. The sessions included from two to four subjects. All subjects present at a given training session heard all of the stimuli.

The subjects took turns answering "same" or "different" after the presentation of a pair of stimuli. The stimuli were divided up so each subject had an equal, or nearly equal, number of stimuli to respond to. When a training group numbered three subjects, one of the subjects responded to only two of the stimuli while the other subjects responded to three. The experimenter told each subject whether the response was correct or incorrect.

All stimuli in this training were selected from the tempi which are available on the Seth Thomas metronome. Those tempi which were excluded were: tempi used on the rhythm test, tempi used in acclimatizations, tempi under 50 beats per minute, and tempi over 200 beats per minute.

One each of the following tempi were used for pairs of stimuli which were alike: 50, 52, 54, 58, 69, 76, 84, 92, 108, 116, 126, 144, 160, 168, 176, and 192. The above tempi were divided into four groups. The slowest four tempi were group one, and so on. One tempo from each group was randomly assigned to one of the four training sessions for Training A.

The selection of tempi for pairs of training stimuli which were to be different was a complicated process because the researcher had

used so many of the tempi in previous work. The tempi were to be unlike any used on the test or in acclimatization. Again, the experimenter excluded tempi under 50 and over 200. The second stimuli of each pair was different from the first by 24 beats per minute as was the case with items in the rhythm test. The following were selected to be the first stimulus of pairs having different rates: 50, 52, 76, 84, 92, 108, 116, 144, 168, and 192. The following are the first of a pair which were followed by a faster stimulus: 52, 84, 92, 144, 168. The following are the first of the pair which were followed by a slower rate: 76, 108, 116, 168, 192. This yields only 10 pairs and 16 were needed. In order to have the necessary number the researcher included the following tempi twice in the training sessions: 76, 84, 92, 108, 116, and 144. These were selected because they are the more moderate tempi. Four of the 16 tempi were randomly selected for each of the four training sessions; however, none of the tempi which was used twice was used more than once in one session.

Each separate stimulus consisted of eight ticks from the metronome. The experimenter slid the rate indicator back and forth on the metronome after the first stimulus and then set the indicator for the second stimulus. This was done even if a pair of stimuli were the same. This procedure was necessary to insure that the subjects did not learn that an unchanged indicator meant an unchanged stimulus. The following are the rates used in each session.

Session I

1. 192-168
2. 92

3. 76-52
4. 92-116
5. 126
6. 144-168
7. 52
8. 168

Session II

1. 108-84
2. 58
3. 52-76
4. 168-144
5. 84-108
6. 160
7. 76
8. 116

Session III

1. 108
2. 176
3. 116-92
4. 168-192
5. 76-52
6. 69
7. 50
8. 108-84

Session IV

1. 84-108
2. 84
3. 144
4. 54
5. 144-168
6. 116-92
7. 92-116
8. 192

Training B

The same tempi which were used for the like stimuli of Training A were used in Training B. The tempi were arranged into four groups. One was selected from each group and randomly assigned to one of the four training sessions. The following procedure was used in training. The researcher set the metronome at the given tempo and allowed for it to

tick about eight times while he clapped synchronously with the ticking. Then, the researcher invited all of the subjects to clap in time with the metronome for about 20 more ticks. Next, the subjects and the experimenter walked around the room for about 20 steps in time with the ticking. The walking was substituted for marching in place during the second and fourth sessions. Finally, each subject, one at a time, was allowed to strike a percussion instrument in time with the metronome for about 20 ticks. Only one instrument was used throughout each of the sessions. The instruments used were snare drum, xylophone, claves, and tom-tom.

After the entire procedure was completed, the same procedure was used for the next tempo. The following are the tempi and instruments used in Training B.

Session I

(snare drum)
1. 126
2. 168
3. 50
4. 76

Session II

(xylophone)
1. 176
2. 108
3. 58
4. 84

Session III

(claves)
1. 144
2. 92
3. 160
4. 54

Session IV

(tom-tom)
1. 69
2. 116
3. 52
4. 192

Training C

The experimenter devised 20 rhythm patterns which are similar to but not exactly like those contained in Test C of the rhythm test. Four patterns each of three-, four-, five-, six-, and seven-note patterns were constructed.

One pattern of each length was randomly assigned to each of the training sessions. The procedure for a session was as follows. The experimenter executed a three-note pattern by either clapping, hitting one hand on a table, clapping both hands above the head, hitting a knee with one hand, or hitting both hands on the floor while seated. All of the subjects were then invited to respond similarly. The researcher executed the same pattern again after which one subject was allowed to respond by himself. This was done over and over until all the subjects had been given an opportunity to respond to that pattern. Next, a pattern of four notes was presented, and the same procedure was used again. This pattern was followed with the five-, six-, and seven-note patterns.

The tempo for each pattern was set by the researcher at about 100-120 beats per minute. If a subject got the correct number of taps or the correct number of taps and correct rhythm, he was told that he had done a good job.

The researcher decided to use the various methods of executing the patterns to add variety to the procedure. All five methods were

used once during each session, and their order was randomized. The patterns which were used in Training C will be presented below. A modified form of musical notation is used. The following figures will be used: o/ (half note), x/ (quarter note), x⌐ (eighth note), x⌐= (sixteenth note).

Session I

1. hands on floor — 6/8 x/. x/ x⌐
2. hit knee — 3/4 x⌐= x/ x/ x/
3. hands above head — 3/4 x/ x⌐= x⌐= x⌐=
4. clap — 5/4 x⌐ x/ x/ x/ x/ x/
5. hand on table — 5/4 x⌐= x⌐= x/ x/ x/

Session II

1. hands on floor — 6/8 x/. x⌐ x/
2. hand on table — 3/4 x⌐ x/ x⌐ x/.
3. clap — 4/4 x/ x/ x⌐= x/
4. hands above head — 3/4 x⌐ x/ x/ x⌐= x/
5. hit knee — 5/4 x/ x⌐= x/ x⌐= x/

Session III

1. hand on table — 2/4 x/ x⌐=
2. clap — 2/4 x⌐= x⌐=

3. hands on floor $\frac{3}{4}$ x. x x x x

4. hit knee $\frac{5}{4}$ x. x x x x x

5. hands above head $\frac{4}{4}$ x x x x x x x

Session IV

1. hand on table $\frac{2}{4}$ x x. x

2. clap $\frac{2}{4}$ $\overset{3}{\text{x x x}}$ x

3. hands on floor $\frac{4}{4}$ x. x x x x

4. hit knee $\frac{4}{4}$ x x x x x. x

5. hands above head $\frac{4}{4}$ x x x x $\overset{3}{\text{x x x}}$

Training D

The researcher devised 20 rhythm patterns for use in Training D. There were four each of three-, four-, five-, six-, and seven-note patterns. These were similar but not exactly like the rhythm test, the acclimatizations, or rhythms in Training C. One pattern from each group was assigned to one of the four training sessions.

Half of the patterns with the same number of notes were randomly selected to be played and repeated unchanged. The other half were played in their original form and followed by a modified form. The modified form was actually one of the similar patterns which was being used in another training session of Training D.

All patterns were played by the researcher at a moderate tempo. A different instrument was used for each training session. The four

instruments used were the tom-tom, claves, xylophone, and snare drum. The subjects took turns telling whether the stimuli were the same or different.

The researcher began with the three-note patterns, and so on, until the seven-note pattern had been played and responded to by one of the subjects. Each subject was told whether his response was right or wrong. The researcher ran through this whole sequence once again. This makes a total of 10 items for each session. The subjects were allowed to respond to the stimuli alternately, and the items were as equally divided among or between the subjects. The patterns which were used in Training D follow below.

Session I

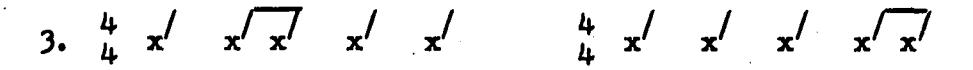
(tom-tom)

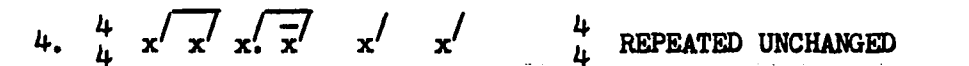
1.	$\frac{6}{8}$ x x x	$\frac{6}{8}$ x x x
2.	$\frac{3}{4}$ x x x x	$\frac{3}{4}$ x x x x
3.	$\frac{4}{4}$ x x x x x	$\frac{4}{4}$ x x x x x
4.	$\frac{3}{4}$ x x x $\overset{3}{x}$ x	$\frac{3}{4}$ REPEATED UNCHANGED
5.	$\frac{4}{4}$ x x x x x x	$\frac{4}{4}$ REPEATED UNCHANGED

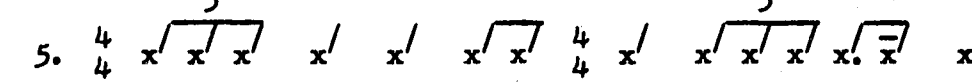
Session II

(claves)

1.	$\frac{2}{4}$ x x	$\frac{2}{4}$ REPEATED UNCHANGED
2.	$\frac{3}{4}$ x x x x	$\frac{3}{4}$ REPEATED UNCHANGED

3. 

4. 

5. 

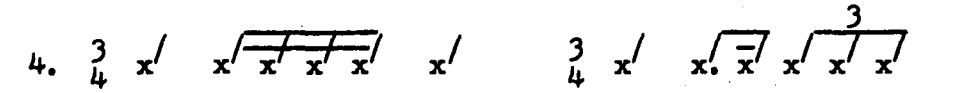
Session III

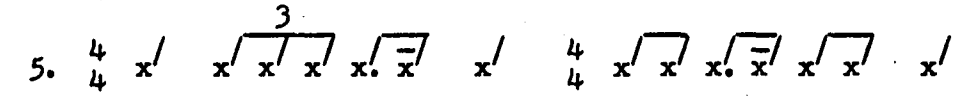
(xylophone)

1. 

2. 

3. 

4. 

5. 

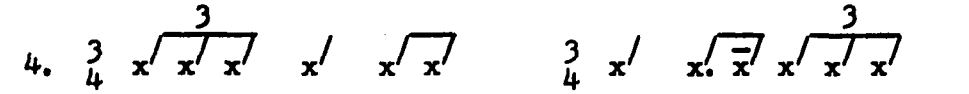
Session IV

(snare drum)

1. 

2. 

3. 

4. 

5.