



1
2000

This is to certify that the
thesis entitled
THE RELATIONSHIP BETWEEN SINGING ACHIEVEMENT
AND TONAL MUSIC APTITUDE

presented by

Sheri E. Jaffurs

has been accepted towards fulfillment
of the requirements for

Master degree in Music Education

Cynthia Taggart
Major professor

Date 5/1/00

PLACE IN RETURN BOX to remove this checkout from your record.

TO AVOID FINES return on or before date due.

MAY BE RECALLED with earlier due date if requested.

DATE DUE	DATE DUE	DATE DUE
370605 MAY 04 2007		

PLACE IN RETURN BOX

to remove this checkout from your record.

TO AVOID FINES return on or before date due.

MAY BE RECALLED with earlier due date if requested.

DATE DUE	DATE DUE
MAR 17 2005	
10 22 04	

THE RELATIONSHIP BETWEEN SINGING ACHIEVEMENT
AND TONAL MUSIC APTITUDE

By

Sheri E. Jaffurs

A THESIS

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

MASTER OF MUSIC EDUCATION

School of Music

2000

ABSTRACT

THE RELATIONSHIP BETWEEN SINGING ACHIEVEMENT AND TONAL MUSIC APTITUDE

By

Sheri E. Jaffurs

The purpose of this study is to develop an understanding of the relationship between tonal aptitude and singing achievement of students. The problem of this study is as follows: is singing achievement as measured by Rutkowski's Singing Voice Development Measure (SVDM) related to tonal music aptitude as measured by Gordon's Intermediate Measures of Music Audiation (IMMA)?

Fifty-six fourth grade students from an elementary school in Michigan were involved in this study. IMMA was given to determine tonal music aptitude. The students were rated individually using SVDM to determine singing achievement. Judges rated the student's singing, and inter-judge reliability was found to be acceptable. Means, standard deviations, and standard errors of measurement were calculated for tonal aptitude scores and singing accuracy scores. These were distributed normally. Split halves reliability for the tonal aptitude scores compared to the norms. Pearson r was used to determine that a significant relationship ($p < .01$) exists between singing as measured by SVDM and tonal aptitude as measured by IMMA. This suggests that singing achievement is related to tonal aptitude.

Copyright by
SHERI ELIZABETH JAFFURS
2000

DEDICATION

In memory of my father, A.W. Luster.

As a child, he woke me for school every morning with these words:

“You’ve got an opportunity to go to school.”

ACKNOWLEDGEMENTS

My appreciation to Dr. Cynthia Taggart, committee chairperson, for her continual guidance and encouragement throughout my degree program and especially with the writing of my thesis.

Thank you to the members of my committee, Dr. Colleen Conway, and Dr. Jonathan Reed for your enthusiasm. To Dr. Joanne Rutkowski, Pennsylvania State University, thank you for your interest and ready responses to my e-mails.

Without the patience and sacrifice of my family this degree would not have been possible. John, Nicholas and Alexis, you are the BITW.

TABLE OF CONTENTS

LIST OF TABLES.....	viii
LIST OF FIGURES	ix
CHAPTER 1	
INTRODUCTION	1
Singing Achievement and Music Education.....	1
Music Aptitude	3
Singing Achievement and Music Aptitude.....	5
Summary.....	6
Purpose	7
Problem.....	7
CHAPTER 2	
RELATED LITERATURE AND RESEARCH	8
Music Aptitude and Achievement	8
Singing Achievement.....	8
Measurement of Singing Voice Development.....	25
Relationship Between Music Aptitude and Singing Achievement.....	31
Summary.....	34
CHAPTER 3	
DESIGN AND METHODOLOGY	36
Participants	36
Design.....	36
Procedure	39
Teaching Procedure for the Criterion Song.....	41
Design and Analysis	44
CHAPTER 4	
ANALYSIS AND INTERPRETATION OF DATA.....	45
Introduction.....	45
Results.....	45
Interpretations.....	48
CHAPTER 5	
SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS	49
Summary.....	49
Conclusions.....	50
Recommendations.....	51
Implications	52
APPENDICES	
Singing Voice Development Measure for Judges	53
Letter to Parent for Permission to Participate in Study	54

LIST OF REFERENCES.....	55
-------------------------	----

LIST OF TABLES

Table	Page
1. Means, Standard Deviations, and Standard Errors of Measurement for the Tonal Subtest of Intermediate Measures of Music Audiation	45
2. Means, Standard Deviations and Standard Errors of Measurement for Singing Voice Development Measure	45
3. Interjudge Reliability for Singing Accuracy Scores With and Without Text and for Singing Composite Scores	46
4. Correlations between Scores for Singing Without Text, with Text, and Singing Composite, and Scores for Tonal Aptitude.....	46

LIST OF FIGURES

Figure.....	Page
1. Aptitude tests created by Edwin E. Gordon.....	4
2. Goetze et al. singing research review	11
3. Rutkowski (1990) pilot study of singing voice classification levels	28
4. Rutkowski (1990) revised singing voice classification levels.....	29
5. Rutkowski (1986) Singing Voice Development Measure.....	38
6. Criterion song for singing achievement assessment.....	39

CHAPTER I

INTRODUCTION

Singing Achievement and Music Education

“Since singing is so good a thing, I wish all men would learn to sing”
William Byrd (Swears, 1985, p. 1).

In our society a certain degree of musical achievement is expected. “We are expected to dance at weddings, cheer at sporting events while clapping hands in time with the crowd, sing “Happy Birthday” to friends and relatives, or share a lullaby with an infant” (Feierabend, 1990, p. 16). For those who find these tasks daunting, it may seem that they lack the ability to sing or achieve in music. Jersild and Bienstock (1934) stated that children become resigned to being non-singers and will “remain in a situation analogous to that of a non-swimmer who has developed the habit of standing on the bank while others leap in the water and try to swim” (p. 497). That this has been a persistent problem is troublesome when we consider the efforts of music educators to help all children develop musical skills.

Historically, there has been a vocal tradition in the elementary music curriculum. “Public school music is rooted in a vocal tradition, and to this day, singing is regarded as a core skill in elementary school music education” (Apfelstadt, 1988, p. 27). From the nineteenth through early twentieth century music educators have emphasized “learning to sing well” (Phillips, 1992, p. 8). Early music study was in the form of singing schools that were the “vehicle for the first formal music education....persons of all ages, who were considered to have musical talents, usually attended these schools” (Rutkowski, 1985, p. 1).

There are many people today who believe that if they cannot sing they are not musically intelligent. Often, adults make comments such as, “I can’t carry a tune; I know nothing about music.” However, it is possible that the ability to sing in tune is not a “fail safe” indicator of overall music aptitude or achievement. Phillips (1992) stated, “how sad the adult who equates the inability to sing with the lack of music ability in general” (p. 3).

Terms such as “tone deaf” and “monotone” imply more than just being a “non-singer.” These terms signify a person who cannot sing because they were not born with musical potential or that they are “a nonsinger because of innate reasons” (Gordon, 1971, p. 93). The perception follows that a student who does not learn how to sing is not just a “non-singer” but also not a musician.

The term “monotone” has had various meanings and can describe various types of problem singers. Many children are categorized as “monotones” when they can vocally produce more than one pitch. Joyner refers to monotones as “someone who consistently fails to reproduce the tonal configuration of a melody in a recognizable manner” (1969, p. 115). The term monotone has been used to refer to singers who: are unable to vocally produce more than one pitch, vocally produce a limited range or pitches, have pitch discrimination problems, are unable to establish tonality, or can establish tonality but are unable to maintain it (Welch, 1979). Gordon (1997) has classified these uncertain singers as the non-singer and the out-of-tune singer. Non-singers have not found their singing voices and out-of-tune singers have found their singing voices but are not able to match pitch because of an inability to sing in the established tonality.

A common assumption is that singing is an innate talent rather than a trainable skill. However, texts, studies, and articles exist regarding the training of singers by authors

such as Apfelstadt (1988) who reported elements needed for singing. Apfelstadt broke the task of singing into perception, remembering and reproducing. Goetze & Horii (1989) established that students sing more accurately when singing individually than in a group. Goetze, Cooper, & Brown (1990) reported the findings of research on singing from 1965 to 1990. Kenneth Phillips' (1992) text is dedicated to Lowell Mason, "the father of singing among the children" (p.vi) and is a vocal pedagogy source.

Music Aptitude

The music aptitude of a child is an important guide for teachers and a "measure of a child's potential to learn music" (Gordon, 1997). According to Gordon, (1971) the music aptitude level of a child is an indication of the potential that child has for learning to audiate. All children are born with some degree of music aptitude. However, the environment shapes aptitude when children are young. Audiation is how one learns to achieve in music; one cannot learn to achieve in music unless one audiates. "Audiation takes place when one hears and comprehends music silently when the sound of the music is no longer or never has been physically present" (Gordon, 1997, p. 19).

There are two types of music aptitude, developmental music aptitude and stabilized music aptitude. Students' aptitude is developmental or stabilized depending on their age. A child is in the developmental music aptitude stage from birth to approximately nine years of age. Developmental music aptitude is "a product of both innate potential and early environmental influences" (Gordon, 1997, p. 10). Children need to be exposed to music at an early age in order to reach the full potential of their innate music aptitude. Stabilized music aptitude is music potential that can no longer be affected by the

environment. Children of approximately nine years of age develop a level of stabilized music aptitude that remains the same for the rest of their lives.

Tests can be given to discover the level of a child's music aptitude. Gordon's standardized aptitude tests are the only standardized tests on the market today for the measurement of music aptitude. A list of these tests are as follows: Audie (Gordon, 1989), which measures developmental aptitude of children ages three and four; Primary Measures of Music Audiation (Gordon, 1979), which measures developmental aptitude of children from kindergarten through grade three; Advanced Measures of Music Audiation (Gordon, 1989), which measures stabilized aptitude of children in grade seven through twelve; and Musical Aptitude Profile, which measures stabilized aptitude of children in grade four through twelve (see Figure 1). Also, Intermediate Measures of Music Audiation (IMMA), (Gordon, 1982) can be administered to ascertain a child's level of stabilized music aptitude. IMMA measures stabilized music aptitude of children in grade four and the developmental music aptitude of children in grade one through grade three.

Primary Measures of Music Audiation	Intermediate Measures of Music Audiation	Audie	Advanced Measures of Music Audiation	Musical Aptitude Profile
1979	1982	1989	1989	1995
Kindergarten to Grade 3	Grade 1 to 4	3-4 years of age	Grade 7 to 12	Grade 4 To 12
Developmental Aptitude	Developmental & Stabilized Aptitude	Developmental Aptitude	Stabilized Aptitude	Stabilized Aptitude

Figure 1. Aptitude tests created by Edwin E. Gordon.

There are many reasons for music educators to administer these tests. Two important reasons are 1) to determine students' global music aptitude and 2) to determine music students' strengths and weaknesses. Teachers may use test results to compare the students idiographically and normatively. An aptitude test provides a teacher with

information to teach to the strengths and weaknesses of each student. They also provide a teacher with the means of identifying a student who would benefit from additional instruction. A complete aptitude test may consist of many subsets of musical aptitudes. There are two main subsets to most of Gordon's music aptitude tests, tonal and rhythm.

Singing Achievement and Music Aptitude

Ball (1995) recognizes two dimensions of music intelligence, music achievement and music aptitude. Music achievements are those skills that can be learned, such as singing or playing the clarinet. Music aptitude is the potential to learn. Elementary teachers who evaluate a child's musical abilities should assess both music aptitude and music achievement. "Only after objective measurement has been made of a child's aptitude and achievement in music can a teacher expect to make a dependable subjective evaluation"(Gordon, 1997, p. 9).

Gordon (1971) stated that the study of music aptitude has been of more interest to researchers than the study of musical achievement. While the difference between aptitude and achievement is important, the two concepts can be linked. Gordon believes that one who possesses high aptitude may not have a high achievement but one who possesses "a high level of music achievement must also have a high level of music aptitude" (1997, p. 42). Ball concurs and states, "the musical outcome is always influenced by basic aptitude" (1995, p. 19). Gordon has consistently found that developmental aptitude scores decline in children who are not actively involved with singing and moving to music.

In 1996, Rutkowski suggested that the "relationship between a child's ability to use singing voice and his or her music aptitude seems to be small" (p. 354). She suggests that students labeled as "nonsingers" may have high music aptitude. There are many possible

relationships between singing achievement and aptitude. A student could have a low music aptitude and low singing achievement. A student could have high singing achievement and high aptitude. Another relationship that Rutkowski (1996) believes exists is high aptitude and low singing achievement.

To establish a relationship, the aptitude and singing achievement of each student must be measured. Rutkowski suggests that the level of vocal skill could denote the potential of that student. The means of evaluating music aptitude have been established. Considering the importance of singing as a core of the elementary program it would be beneficial for educators to develop tools to measure achievement accurately so that inaccurate singers could be identified to receive remedial training.

Summary

Does singing achievement denote tonal potential? Listening to the individual singing of students to determine singing ability is a daunting task to many music educators. Music specialists are often faced with large numbers of students each week, and it is difficult to listen to these students individually. Even if this task could be accomplished, music educators may be reluctant to ask students to sing by themselves.

Gordon (1971, 1997) believes that the development of the singing voice and the extent to which a child is able to “express themselves musically with their singing voice”(1997, p. 252) is dependent upon the aptitude of the child. However, Rutkowski (1996) has found that singing achievement may or may not be related to aptitude.

Teachers may be guilty of assuming that singing ability is an indication of tonal aptitude or vice versa. If a student cannot learn to musically achieve at a high level,

specifically in singing, without high music aptitude, then does it not follow that a student with high music aptitude should have a high degree of singing achievement?

Defining the relationship should improve knowledge of how children learn to sing. It may also give music educators insight into how to teach students, if they exist, who have high music aptitude but are nonsingers.

Purpose

The purpose of this research is to gain information about the relationship between singing achievement and tonal music aptitude to improve singing instruction and to deepen the music education profession's understanding of music aptitude.

Problem

The problem of this study is to determine whether singing achievement as measured by Rutkowski's Singing Voice Development Measure is related to a child's tonal music aptitude as measured by Gordon's Intermediate Measure of Music Audiation. The following are definitions within the context of this study:

Achievement: what a child has already learned.

Aptitude: the potential of a student.

Monotone singer: a singer who consistently produces incorrect pitch or pitches when trying to reproduce a pitch or melody.

Non-singer: the child who may not be able to do one or all of the following:
sing in head voice, sing in-tune, or sing a range of pitches.

CHAPTER II

RELATED LITERATURE AND RESEARCH

Music Aptitude and Achievement

A child's level of music achievement should be at least as high as his or her aptitude but a child with high music aptitude may not display high music achievement. "Regardless of the level of music aptitude one is born with, unless early informal influences in music ... are favorable, that level will never be realized in achievement" (Gordon, 1986, 1979).

Singing Achievement

Singing in Music Education

In 1838 Lowell Mason advocated that all children could learn to sing. "Mason knew then what many of today's music educators have learned—singing is a skill that can be taught..." (Phillips & Aitchison, 1997, p. 185). Welch (1979) concluded that poor singing is a product of the environment and not an innate trait. Singing, according to Welch, can be taught and "pitch discrimination is a function of vocalization" (p. 56). Many prominent music educators support the idea that singing is a learned behavior (Levinowitz, Barnes, Buerrini, Clement, D'April, & Morey, 1998; Gordon, 1997; Rutkowski, 1996; Phillip, 1992; Atterbury & Silcox 1991; Goetze & Horii, 1989; Swears, 1985).

This early directive to music teachers from Lowell Mason cultivated the move toward a singing-based approach to music education. Phillips and Aitchison (1997) noted that for a hundred years the development of singing was a strong part of the elementary curricula. The Italian bel-canto method employed exercises in posture, diction, tone

formation and breathing. Around the turn of the century, a music education movement away from this approach began. Worried that students were not experiencing joy in making music, music educators advocated the singing of songs rather than vocal exercises.

This “singing songs” approach has been under scrutiny by today’s music educators. Phillips and Aitchison (1997) cite numerous studies that show positive effects from instruction on singing achievement. Their study was a replication of an earlier study in 1983. The 1997 study showed a positive effect on the singing achievement of students who had been taught formal voice training addressing quality of singing voice, pitch perception, breath capacity, and shallow breathing. Breath support, vocal range, and pitch accuracy improved. The researchers determined that those who were taught using formal vocal techniques used their breath better than those who were taught using a traditional singing-song approach.

Jersild and Bienstock (1934) completed two studies with four hundred children, ages three to eight. The first study examined “the effects of training on children’s ability to sing tones, intervals and songs” (p. 494). The results of the study were used to compare the number of tones a child could sing at various ages. It was determined that a child reaches the extent of their vocal range by the third grade. In the second study, thirty-six children were divided into two groups for an investigation. The study examined the effect of vocal training, forty ten-minute periods, over six months. There was a significant gain in the number of pitches the subjects were able to sing, as compared to the control group that received no training. The authors found that children who were taught to sing

showed a significant improvement in singing achievement over children who had no instruction.

Welch (1979) reviewed the research regarding “poor pitch” singing from 1937 to 1972. He identified twelve possible causes of “poor pitch” singers. There are five major categories of reasons for “poor pitch” singing: perceptual difficulties, social, behavioral, technical, and physiological reasons. Pitch perception problems were believed to be from an inability to discriminate and remember tones. Welch concluded that the majority of the research suggests that “pitch discrimination is a function of vocalization” (p. 56).

Apfelstadt (1988) reports that, while heredity and environment may be important to singing achievement, there are three main components of vocal accuracy: hearing, remembering, and reproducing. She believes that awareness of these components and the means to acquire them must remain a top priority for music educators.

Review of Studies Related to Singing Achievement

According to Goetze & Horii (1989) singing ability has been of persistent concern to elementary music teachers for the past sixty years. The majority of researchers who have examined singing have looked at a single dimension of vocal production. The experiments that have been done seem to confirm that, by isolating one factor of singing, we can prove or disprove its importance in vocal production.

Goetze, Cooper, and Brown (1990) conducted a literature review of the current research on singing in the general music classroom. Studies of singing were categorized by description, processes, and accuracy factors. These studies give a good overall view of the research topics relevant to singing ability and the findings of the reports (see Figure 2).

<u>RESEARCH TOPIC RELATIONSHIP</u>	<u>GENERAL FINDINGS</u>
Age to singing ability	Positive relationship
Pitch discrimination ability to vocal accuracy	Inconclusive
Vocal range, register to pitch and vocal model	May be affected
Melodic direction to singing accuracy	Inconclusive
Harmonic accompaniment/a cappella to singing accuracy	Inconclusive
Timbre and pitch of the vocal model to singing accuracy	May be affected
Singing in a group to singing individually	More formal research needed
Singing with text to singing without text	Inconclusive
Teaching breath control technique to singing ability	Positive relationship
Verbal and visual feedback instruction to singing ability	Positive relationship
Vocal instruction to inaccurate singers	Positive relationship

Figure 2. Goetze et al. Singing research review and the general findings of each study.

Individual Versus Group Singing

Researchers from the last ten years have looked at accuracy of individual singing as compared to accuracy when singing with a group. Greenberg (1979) encouraged teachers and parents of young children to promote individual singing. Gordon (1971) stated that students who sing accurately when alone and are unable to sing accurately in a group have a higher level of music aptitude. According to Goetze & Horii (1989) individual singing may not have been the central issue in many studies, but it has been

observed and reported since as early as 1938. They warned educators that classifying singers as inaccurate if an assessment is made from group singing is not appropriate. They note that there is “little evidence... that group singing is the most effective setting in which children learn to sing accurately” (p. 57). Their 1989 study concluded that there was a difference in the singing accuracy of students who sing individually and students who sing in a group. The fact that so many investigators mentioned this phenomenon led Goetze to the purpose of her study.

Goetze examined one hundred kindergarten, first, second and third grade students to determine if they would sing more accurately when singing, both with and without text, individually as compared to when singing with a group. Goetze tested six subjects at a time; five subjects observed and reviewed the song as one student was asked to individually echo a task phrase of two measures, twice, once with text and once without text. Then all six subjects were asked to sing a six-measure melody in unison. Goetze found that singing accuracy was better when students sang individually as compared to when they sang with a group.

A Visi-Pitch was used to determine pitch accuracy when singing in unison, individually, with text and using a neutral syllable. Goetze also found that children sing more accurately when using the syllable “loo” instead of text. She reported that students made significant improvements in pitch accuracy between kindergarten and first grade, and in shape, or contour, they made improvements between first and third grade. Phillips (1989) supported these findings as reliable and relevant to music educators. He listed the strengths of the study as being the use of the Visi-Pitch, which analyzes pitch accuracy objectively, and clear and extensive data analysis. Phillips questioned the conclusion

Goetze reached regarding the improvement students made in pitch accuracy and contour determination. He reports that Goetze concluded and discussed these factors without placing them in the final results of the study. If Goetze's assertion is correct, then discrepancies exist between these findings and prior research. Although the study was not generalizable because students were not chosen randomly, its findings are significant and justify further research in this area.

Cooper (1995) used a Visi-Pitch with computer interface to determine the singing accuracy of one hundred sixty-nine first through fifth grade students. The research problem focused on the relationship between individual versus group singing and singing accuracy. Children were asked to echo a four-beat melodic pattern that was recorded by a child-voice model. The subjects were asked to sing individually. Then they were asked to sing a pattern in unison with a recorded vocal model on a tape. There was no significant relationship found between individual versus unison singing and tonal accuracy.

Singing with a Text and Singing with a Neutral Syllable

Goetze (1985) reported that singing with text combines the physiological processes necessary to form words and the cognitive processes necessary for language. Further, she states that it has been established that language is a left hemisphere function and research suggests that singing is a right hemisphere activity. Therefore, if true, different hemispheric activities may be the reason for the difficulty some children demonstrate in singing songs with a text.

Levinowitz (1987) conducted a study to gather information regarding the use of text and no text in song instruction. Three classes of kindergarten children and three classes of first grade children participated in the study. The children were given PMMA

as a pretest before instruction began. Three types of singing instruction were used for one year. Children in experimental group one sang most of the songs with words. Children in experimental group two sang songs primarily without words. Children in experimental group three sang all songs with words. After the instruction period, PMMA was re-administered. The children were individually given a singing achievement test that was tape-recorded. After analysis, Levinowitz concluded that audiation of kindergarten and first grade children with low developmental music aptitude improves with song instruction both with and without text. There was no indication that singing instruction with or without text has an effect on children with high developmental music aptitude. There was no significant difference between the singing achievement of the entire group of children who sang with text and those who sang without text.

Levinowitz (1989) analyzed the singing achievement of preschool children, ages four and five. She investigated the relationship between singing intonation and children's language development using songs with and without text. Levinowitz analyzed the children's rhythm and tonal singing achievement by developing two criterion songs. The first criterion song had text and the second criterion song, with a new melody, had no text. Using a continuous rating scale, two judges rated the children's performances of the criterion songs. The Peabody's Picture Vocabulary Test (Dunn, 1959) was the instrument used to measure the children's language development. She concluded that there was no relationship between language development and rhythm or tonal achievement as measured by performing rote songs that are sung with or without text.

Lange (1999) investigated the effects of using text in the performance of a song. Fifty-eight kindergarten students from four classes of two elementary schools,

participated in the study. PMMA was administered prior to instruction. One group received music instruction that used text when performing songs. The other group received music instruction that used a neutral syllable. After a 24-week instructional period, PMMA was re-administered to students. In addition, students were asked to sing a familiar song using text and were rated on their tonal accuracy. Lange states that students with low aptitude scores who received instruction using neutral syllables had the greatest increase in tonal aptitude. Lange reported that there was not a difference between the text and no text groups after the instruction period in the performance of resting tone.

Vocal Models

When testing singing accuracy, the researcher must decide which type of model will be used to demonstrate the pattern or song. Singing accuracy researchers have investigated how different models affect vocal accuracy. Studies examining the effects of this variable are numerous. Researchers such as Rupp (1994), Green (1989), and Greenburg (1979) investigated which mode would be easiest for the subject to emulate. Some of the modes examined have been the following: instrumental models, piano, oscillators, female and male adult models, and peer models. Researchers who have examined the teaching vocal model for children suggest that teachers who model singing by using head voice achieve higher vocal accuracy with students than those who do not (Jones, 1994). Green (1989) examined three sound sources to determine timbre preference for the following vocal models: 1) child model, 2) female model, and 3) male model. She found that the peer, child model is the best in helping children sing accurately, followed by a female model and then a male model.

Kramer (1985) examined the effects of two different types of singing instruction. Two hundred one third and fourth grade students participated in the ten-week study. One group received a traditional approach toward singing instruction and the second group received instruction using Gould's Specialized Program in Singing (1969). Kramer wanted to establish that this "creative teaching method which incorporates auditory, kinesthetic images, visual mediating strategies, and mnemonic association" (Kramer, p. 19) would improve singing instruction.

Kramer's initial singing test was administered with a male voice singing the examples. The data from the first test indicated that 50% of the students were inaccurate singers. A second singing test was administered with a male voice singing with a falsetto voice. The number of inaccurate singers from the second test was reduced to 30%. Kramer claimed that the male vocal model was confusing to the third and fourth grade students. He stated that the "male singer, and the tonal timbre of a mature female, may be causing great problems" (p. 111).

Yarbrough, Green, Benson, & Bowers (1991) examined one hundred sixty-three students in a university laboratory. The subjects were chosen based on their inability to respond to a male and female model. Yarbrough wanted to study which factors affected an inaccurate singer's ability to match pitch. The subjects were from kindergarten through eighth grade. The students were instructed using the Kodaly solfege and hand-sign system for two months. After the instruction period, subjects were randomly divided into three different response groups. One third of the students sang using hand signs alone, one third of the students sang using solfeggio syllables alone, and one third sang using the syllable "la." The subjects were tested twice, once with the female model and

once with the male model. Subjects who had already been identified as inaccurate singers were tested to determine what factors would affect their ability to match pitch. There was a significant difference between responses to male and responses to female with responses to the female model being more correct. Although hand signs seemed to help the students sing more accurately than the other two modes, no statistical difference was found. The researchers speculated that there could be a relationship between response mode and age level. Older students respond better to hand signs than do younger students.

The model a child is asked to match has been of great interest, especially to male teachers. Goetze et. al (1990) speculates that the natural inclination a child has for imitating a model can be of influence to the singing production of a child. "The presence of a model pitch stimulus and the quality and pitch of that model can inhibit or enhance a child's ability to sing accurately" (p. 23).

Singing Achievement and Piano Accompaniment

Atterbury and Silcox (1993) report that "the development of the ability to sing does not occur similarly in all young children" (p. 40). After identifying previous research involving remedial training of inaccurate singers, they wanted to investigate some of the causes of poor singing. They noted that singing is a coordination of reproducing what one has heard and remembered. Their year-long study examined the effect of the use of harmonic piano accompaniment on the singing achievement of kindergarten students.

Two hundred five kindergarten student participated in the study. One group of students were taught with piano harmonic accompaniment and the other group had no accompaniment during instruction.

Atterbury and Silcox found no significant difference between the singing achievement of kindergarten students who had received instruction with piano harmonic accompaniment and those who did not receive instruction with piano harmonic accompaniment. The posttest singing scores were compared to the aptitude scores of students. The posttest scores of students with high aptitude had higher singing achievement than the posttest scores of the average or low aptitude students' singing achievement. This suggests that aptitude is a more potent factor than use of accompaniment when determining singing achievement.

Use of Hand Signs

Youngson and Persellin (1999) investigated the singing accuracy of first grade students (n=52) who were taught using solfege and Curwen hand signs as compared to those who were taught using solfege without Curwen hand signs. The subjects were given a ten-week treatment. Youngson and Persellin found no significant difference between the post-test scores of the treatment groups. However, the group who received Curwen hand sign instruction did score significantly higher on the posttest than on the pretest for vocal accuracy.

Relationship Between Singing Accuracy and Pitch Perception

Perception is defined by Webster as “the act, process, or result of perceiving.” (1984, p. 520). Carlsen (1981) defines perception as “the extraction of information contained in the structure of a stimulus in the perceptual field” (p. 2). Music is an aural art; musicians perceive from stimuli and respond. Carlsen identified 15 perception skills needed by musicians; seven of the fifteen dealt with pitch perception. These seven pitch-related skills are as follows:

- ability to tune an instrument with standard pitch
- ability to perceive pitches in a melody
- ability to perceive melodic pattern
- ability to perceive melodic variations or changes
- ability to perceive more than one melody in a polyphonic selection
- ability to perceive tonality
- ability to vocally blend.

Radocy and Boyle (1997) discussed perception in the context of musical behavior. Musical behaviors are studied by music educators to learn how inborn and environmental factors influence reactions to music stimuli. "Perception is a process of sensing the environment; it obviously is essential for behavior" (p. 4). To study perception is to examine the "evidence of its results" (p. 4).

Deutsch (1982) wrote and edited a series of treatises pertaining to cognition and perception. These writings examine how persons perceive music stimuli. For example, humans perceive rhythm and tempo by its "temporal order" (p. 175). She compares auditory perception to the rapidly changing spectrums of light and patterns our retina is bombarded with and must interpret. The auditory process involves interpretation of complex and rapidly changing acoustical information. The process of perceiving pitch is "widely acknowledged to be an essential part of musical perception" (p. 396) but when persons perceive a melody, they do so as a Gestalt or pattern, not as an individual pitches. Individual pitches and intervallic identification are not necessary for melodic perception. While it may be important in language perception to perceive individual phonemes it is not as important to the perception of melody or "even the production of music" (p. 264-265).

Petzold (1963) described aural perception as tonal and melodic memory. He equated the ability to conceive pitch as a necessary factor in vocal control. The purpose

of his study was to examine the ways in which children perceive and respond to pitch symbols as they begin to read music. Six hundred children from grades one to six were asked to sing a response after listening to a stimulus. The 45-item test is not available from Petzold's study, and it is not clear if the stimulus was a vocal example or something other than singing.

Petzold stated that "there seemed to be no other way of effectively measuring accuracy of aural perception except in terms of some type of musical performance" (p. 27). He also acknowledged that the use of a vocal response may create a problem in that an inaccurate response could be due to some vocal problem and not from poor perception. An elaborate scoring system was developed to try to alleviate this problem. This scoring ranged from a perfect score for an exact replication of the stimulus to the lowest score for a response that "follows the general direction of the stimulus but contains neither the correct pitches or the correct number of tones" (p. 28). Petzold concluded that a relationship exists between the age of a child and auditory perception. He believed that children reached a plateau of auditory perception by age nine. He concluded that aural perception is from intelligent thought and not from mechanical imitation.

The strength of Petzold's study is the examination of musical stimulus and perception. How music is processed is of interest to educators. It would be beneficial to know the stimulus for the testing; if the stimulus was sung and the student immediately imitated the example, then it is possible that imitation, not aural perception, was evaluated.

To a certain degree, all perception is cognitive. Perception can only be studied by results of an observed behavior. Petzold's study found contradicting information regarding tonal perception. His study measured aural perception using musical performance, which he recognized as being a poor measure if the child could perceive tonally but had not learned to sing.

Joyner (1969) studied the capability of the voices of children classified as monotone singers and those children's abilities to perceive pitch and remember tonal patterns. Thirty-two subjects sang the British National Anthem in a series of keys and were classified by number of pitch errors. A more extensive scale was created to establish which singers were true monotones. The four ranks of the scale are: normal singer, able to sing in tune; Grade A monotone, tuneful at low range and untuneful at usual ones; Grade B monotone, irregular at usual pitches and low range; and Grade C monotones who are completely untuneful (p. 117). The next phase of Joyner's study compared monotones to normal singers. Tonal memory, pitch discrimination and vocal reproduction were tested. The scores were analyzed from one hundred thirty four normal singers, twenty-six Grade C monotones, and twenty grade A monotones. The results of the testing revealed that monotones have deficiencies in pitch discrimination and tonal memory when compared to normal singers. Remedial training of the Grade C monotones revealed that as these children were given vocal training, which included teaching them vocal techniques, their vocal development improved. At the same time, their ability to recall material also improved. As the children learned how to use the vocal instrument pitch discrimination and tonal memory improved. This would seem to indicate that pitch

discrimination may not be as big an issue in the perceptual abilities of monotoners, but rather, the ability to sing enables children to accurately reproduce what they perceive.

Ramsey (1983) investigated the effects of singing ability, age, and training using a pitched instrument on the auditory perception of melody. A need to categorize the singing ability of the subjects in the study led to the development of Preschool Singing Ability Level Test (PSALT) (1983). Ramsey piloted a study of PSALT to determine the length of time needed to administer PSALT and its reliability. Twenty participants were chosen randomly from those who would participate in the main study. He reported a reliability coefficient of .71 for PSALT.

Ramsey was interested in how young children of different ages focus on varying aspects of melody. This problem and an interest in the effect of using a pitched instrument on melodic perception were of chief concern to the main study.

One hundred fourteen subjects from four intact preschools participated in the study. Singing, it should be noted, was not used as a response mode for perception but as a variable in the investigation. The twenty subjects were categorized by singing ability using PSALT (Ramsey, 1983).

The study began with familiarizing the children with the testing room. Treatment groups were either instrumental or non-instrumental. The instrumental group used tone bells to reinforce singing; the non-instrumental treatment group echo-sang. After two weeks, the subjects were taped-recorded singing three songs. The researcher transcribed each response to determine "melodic perceptual variables, absolute pitch, melodic contour, melodic rhythm, tonal center, and melodic interval" (p. 137).

Transcription accuracy was verified by two judges. Intra-class correlations between the transcription and scoring were reported as .99. The results suggested that children are able to perceive contour, rhythm, and interval aspects of melody. Ramsey reported a relationship between singing ability and melodic perception as demonstrated through vocal response. There was no evidence that the use of pitched instruments affected music perception.

Apfelstadt (1984) believes that pitch perception development is necessary to the development of vocal accuracy. Her premise is that a relationship exists between auditory perception and vocal accuracy and that perception of melodic contour could be an important component of singing accuracy. Her study investigated the effects of melodic perception instruction on vocal accuracy and pitch discrimination.

Sixty-one kindergarten children participated in the study. There were five parts to the study: “orientation, protesting, instruction, posttesting, and data analysis” (p. 17). There were two treatment groups and one control group. The first treatment stressed melodic perception through vocal instruction that used kinesthetic and visual reinforcements. The second treatment stressed imitation with attention to musical elements. The control group used traditional vocal instruction with a nonconceptual, rote method. The pretest consisted of the administration of Gordon’s Primary Measures of Music Audiation (PMMA), a vocal accuracy test developed by Boardman (1964), and a rote-song test. After an eleven-week instruction period, the students were given a posttest of PMMA, Boardman’s vocal accuracy test, and the rote-song test.

The results of the study suggest that there was no difference among the groups between auditory discrimination of pitch after instruction. The researcher concluded

“melodic perception instruction did not significantly improve pitch discrimination or vocal accuracy on rote-songs” (p. 22). Several factors were cited as possible problems with this study. First, PMMA measures same-different and not directionality. A test of directionality perception may have been more appropriate. The youngest group of subjects made the most improvement, which may be due to developmental readiness. A larger sample of older students was recommended.

In an initial investigation, Feierabend (1984) found that there was a small relationship between first grade students' ability to aurally perceive tonal patterns and perform the same tonal patterns. He speculated that aural perception skills might not develop concurrent to oral skills, such as singing. Further, many tonal patterns from the Kodaly method that children are asked to sing are difficult to perceive aurally. Conversely, tonal patterns Gordon uses because of their aural perception difficulty levels are difficult for children to sing. With this problem established, Feierabend began his study.

Three first grade classes served as the treatment groups and one first grade class served as the control group. Prior to treatment, all the students were given both singing and listening tests. The same tonal patterns used for aural discrimination were used for the singing test to determine the degree of difficulty children would have singing Gordon's tonal patterns. Each experimental group was given singing instruction that included the singing of two-note, three-note, and four-note patterns, without text, and harmonically between tonic and dominant functions. Experimental group one sang patterns that had been identified as easy to perceive aurally, regardless of singing difficulty. Experimental group two sang patterns that were easy to sing regardless of the

level of perceptual difficulty. Group three sang patterns that were easy to sing and easy to aurally perceive. After a daily treatment period of seven weeks the students were given the listening and singing test again. Data revealed that the treatments did not have an effect on singing or aural discrimination abilities. Feierabend recommended further study with a larger variety of levels between each of the variables and a longitudinal study examining aural skills and singing skill development.

It would seem logical to conclude that children who can sing accurately also demonstrate the ability to discriminate pitch accurately. However, Goetze, Cooper, Brown (1990) reported that there were students who could discriminate pitch accurately but were not accurate singers. The weakness of many of these studies has been using a singing test to measure both singing and pitch perception. These tests may show positive relationships because they are essentially measuring the same thing. There may be a difference between production and perception. Conflicting reports suggest that more research is needed. Care should be taken in the measurement tools used to determine pitch perception and singing accuracy.

Measurement of Singing Voice Development

Although there are many researchers who have assessed the singing voices of children, there is a lack of consistency found in the criterion measures used for these assessments. Many researchers have designed their own test, as was the case of Ramsey (1983) who wanted to examine the singing ability of preschool children. Ramsey designed the Preschool Singing Ability Level Test (PSALT). PSALT is a two-part measure. Part one is designed to measure a child's competency in reproducing pitches from B \flat to C." The characters from the story "Good Night, Owl!" (Hutchins, 1972) are

assigned specific sounds that are to be imitated vocally by the child as the story is read.

Part two of PSALT measures a child's ability to sing a song of their choice.

Rutkowski (1990) reviewed the rating scales researchers used to determine vocal accuracy from 1961 through 1990. She found that these scales examined the ability to sing as defined by accuracy of intonation. Rutkowski determined that it is not possible to rate a child's intonation if they have not found their singing voice. She compared this to trying to determine the intonation accuracy of an instrumentalist who is still trying to learn fingerings.

The problems encountered when trying to rate the singing achievement of subjects is evident when examining some of the rating scales and tests. Many singing achievement tests attempt to determine singing achievement by rating a large number of variables and skills. This is possibly because singing has many facets: head voice, chest voice, varying levels of pitch accuracy, intonation, monotone singing, neutral syllable, and singing with text. The extent to which some singing tests try to encompass all of these variables can be found by looking at an example of a singing test, "The Gould Speech and Song Response Test" (Kramer, 1985), which was created in 1966 to determine singing ability and to measure improvement in singing. An overall score for this test is given based on 318 possible points. This test can be used to evaluate students on their ability to demonstrate singing voice and speaking voice, direction of vocal inflection, intonation, use of text and neutral syllable, phrase singing, sustained note singing, and pitch perception. The test requires students to respond by both singing and answering written questions.

The Singing Voice Development Measure (SVDM) (see Appendix A), developed by Rutkowski, is an instrument used to determine the various stages of a child's singing

voice development and has been examined extensively to ascertain its uses and effectiveness. It has been found to be a valid tool for testing the use of singing voice in children. SVDM was published in *The Quarterly Journal of Music Teaching and Learning* in 1990. It is a five point rating scale that ranks children's voice development from those who cannot sing but chant through children who use head voice over an extended range. Rutkowski (1986) states that SVDM is a valid test of children's use of singing voice.

The purpose of the 1984 pilot study was to determine the reliability and validity of the instrument developed by Rutkowski to measure children's singing voice. Thirty-five kindergarten children participated in the study. The song selected for the study was familiar to the children and in harmonic minor. The children were asked to sing the song. The performances were tape-recorded and evaluated by four judges. The judges evaluated the singing using five classification levels (see Figure 3). The judges also were asked to assess the effectiveness of saying "ready sing" before each student was tape-recorded. Thirteen random performances were evaluated for inclusion in testing procedures. The interjudge reliability was .96 without the "ready sing" procedure and .90 with the "ready sing" procedure. The response frequency also confirmed that the judges used all levels of the classification to rank students. Response frequency was evaluated because the exclusion of one or more of the rating scale levels could have indicated that a particular level was unclear or not necessary. The interjudge reliability for response frequency ranged from .84 to .96.

1. Children who use only speaking -voice inflection but do not sustain tones.
2. Children who exhibit use of melodic contour and sustained tones, but use speaking range or a very high range
3. Children who use a very limited singing range, usually D3 to F#3. This stage has been noted by several other researchers (Harkey, 1979, Joyner, 1971; Young, 1971)
4. Children who use initial singing range, usually D3 to A3; and
5. Children who are able to sing over the register lift, Fp3 and above (Gordon, 1971; Smith, 1963; Young, 1971).

Figure 3. Rutkowski (1990) pilot study of singing voice classification levels.

The pilot study indicated that SVDM “appeared to be an appropriate instrument for assessing the use a child has of his or her singing voice” (1990, p. 87). The judges recommended using “ready sing” to begin each test, although, the interjudge reliability compared to the exclusion of “ready sing” was slightly lower. The judges “felt that it greatly assisted with their evaluation of the voices” (Rutkowski, 1990, p. 87). After the 1984 study, the scoring level descriptors were revised to evaluate the singing of tonal patterns instead of the performance of a song. The use of the term “children” was also deleted in favor of the term “singer” (see Figure 4).

1. "Pre-singer" does not sing but chants the song text.
2. "Speaking range singer" sustains tones and exhibits some sensitivity to pitch but remains in the speaking voice range (usually A2 to C3).
3. "Uncertain singer" wavers between speaking and singing voice, uses a limited range when in singing voice (usually up to F3).
4. "Initial range singer" exhibits use of initial singing range (usually D3 to A3).
5. "Singer" exhibits use of extended range (sings beyond the register lift: B ϕ 3 and above).

Figure 4. Rutkowski (1990) revised singing voice classification levels.

The revised SVDM evaluation had been implemented in the 1986 main study. One hundred sixty two kindergarten children participated in the study. All children were given both a pretest and a posttest using SVDM. The pretest preparation consisted of teaching the students a criterion song and five tonal patterns. The children were divided into two groups. Children in group one sang the song, then tonal patterns. Children in group two sang the tonal patterns, then the song. The posttest preparation consisted of a review of three criterion songs. The posttest administration followed the same procedure as the pretest but this time group one sang the song first then the tonal patterns; the order was reversed for group two. Two raters from the pilot study who had the highest interjudge reliability were the judges for the main study. To eliminate rater bias that could result by listening to a second part of the test, either tonal patterns or song, after listening to the first part of the test, either tonal patterns or song, the decision was made to tape songs on one tape and tonal patterns on another. In other words, the judges listened to one recording of tonal patterns then one recording of songs. Internal consistency within the raters was investigated by including one class of students twice on the song tapes, another

class twice on the tonal pattern tape. Only intra-judge stability was calculated. High reliability coefficients were reported for SVDM. A lower coefficient was reported for the interjudge reliability on the tonal patterns for the pretest but it was higher on the posttest. This was attributed to familiarity and experience with the rating process. Standard deviations and standard errors of measurement were comparable for the pilot and main study.

Levinowitz, Barnes, Guerrini, Clement, D'April, and Morey (1998) conducted a study to assess the reliability of Rutkowski's test. One hundred seventy students from grades one through six participated in the study. The students were taught two criterion songs. Six judges rated students' tape-recorded performances. Judges' scores were used to determine inter-judge reliability, which was .98. The authors recommended SVDM as a reliable means for evaluating the singing voice development of students from kindergarten through fifth grade. They suggested that it should be used with a short, familiar, major song. They also found that the test is compromised when using it with older students due to the use of chest voice singing, which can be a common practice among older students.

Youngson and Persellin (1999) extended Rutkowski's SVDM to a 14-point scale to use when rating singers that use their singing voice but are not accurate singers. Their rating scale is the Vocal Accuracy and Voice Development Measure. The 5 points of the SVDM relate to the ability to sing as opposed to chanting or speaking. The additional ratings in Youngson and Persellin's measure extend the measurement beyond singing voice into accuracy of intonation.

Relationship Between Music Aptitude and Singing Achievement

Gordon (1971) stated that anyone (with the exception of those with physical handicaps) could learn how to sing. He notes that "the degree to which good tonal habits are developed, however, is dependent upon tonal aptitude" (p. 93). He suggests that achievement tests should be used to determine if the level of achievement matches the level of the student's aptitude (p. 130).

Rutkowski (1990) developed SVDM based on the hypothesis that "use of singing voice is a separate but requisite behavior to the ability to sing with accurate intonation" (p. 3). Rutkowski used SVDM in a study she conducted from September to June. The results of the 1996 study concurred with result from the 1986 study in which she found the relationship between the use of singing voice and developmental music aptitude, specifically tonal, to be small (1986). The main purpose of the 1996 study was to determine the effectiveness of individual/small group instruction compared to large group instruction. Rutkowski reported that both treatments had an effect on the singing ability of children. The developmental tonal aptitude increased for children in both groups over the September to June period. Rutkowski questioned the limitations of the five level scoring SVDM and a criterion song of four tonal patterns. She suggested the need for revising SVDM and called for further research investigating instruction of singing voice and developmental music aptitude.

In a pilot study, Rutkowski (1996) examined the effect of individual and small group singing lessons on the pitch accuracy of five-year-old children. She also included developmental music aptitude as a variable. Fourteen kindergarten students participated in the pilot study to determine the strength of a relationship between developmental

music aptitude and use of singing voice as determined by SVDM. The feasibility of using SVDM with large numbers of students and the interjudge reliability of SVDM were also investigated. She concluded that the procedures she used for the pilot study could be replicated with large numbers of children. Interrater reliability coefficients of .84 on the pretest, .95 on the midtest, and .93 on the posttest were acceptable. Based on these results, further study was indicated.

In the main study, the tonal section of Gordon's Primary Measure of Music Audiation (PMMA) (1979) was administered with SVDM to ascertain the strength of the relationship between the scores on SVDM and scores on the tonal subtest of PMMA. The pilot study revealed that there was not a significant relationship between the rhythm subtest of PMMA and SVDM; for this reason the rhythm section of PMMA was deleted from the study.

Rutkowski's participants were ninety-nine students from Pennsylvania. These students received small group/large group singing lessons. Prior to treatment the students's teacher administered the tonal section of PMMA and SVDM. The tonal section of PMMA was also administered halfway through treatment. At the end of the treatment period, the students were given tonal PMMA and SVDM. Pearson product-moment correlations revealed a weak relationship of .21 between singing voice and developmental tonal music aptitude after instruction.

Atterbury and Silcox (1993) studied the results of piano harmonic accompaniment on the singing ability of kindergarten students. The students were rated using Rutkowski's SVDM and Gordon's Primary Measures of Music Audiation (Gordon, 1979) which was given in the fall and again in the spring. There was no significant

difference between the middle and low aptitude students in their ability to sing a song. However, a significant difference was reported between the students who had high aptitude and those with middle and low aptitudes. The use of piano accompaniment to improve singing ability revealed no significant differences.

Atterbury and Silcox decreased SVDM from five scoring levels to four, because they felt that one level was redundant. However, after testing they concluded that the four level range was too narrow and noted that this “may be responsible for the lack of difference observed in the statistical analysis” (p. 45). They suggested that further studies might consider expanding the range of SVDM.

Conkling (1994) examined the effects of Gordon’s Learning Sequence Activities and vocal development exercises on seventh grade students. She also included aptitude as a variable. After a training session of fourteen weeks, the students were given the Advanced Measure of Music Audiation (AMMA) (Gordon, 1989) and asked to perform one choral etude and three solos. The performances were tape recorded and rated independently by three judges. The author reported that students attained higher levels of achievement with high aptitude as compared with students who had moderate or low aptitude. Her findings also suggested that the combination of learning sequence activities and vocal training improved students’ vocal achievements.

Mota (1997) conducted a study in Portugal over a three-year period with children from first through the third grade. The study examined the relationship between music aptitude as determined by the results on Gordon’s PMMA and performance of a set of musical tasks. The musical tasks were singing a song of the child’s choice, singing a criterion song, and playing rhythmic patterns on the xylophone and drum. There was no

significant relationship between the scores on PMMA and performance of the music tasks. The data indicated that students who scored high on PMMA did not necessarily score well on the two singing tests. Conversely, there were some students who did well on the singing tests who scored low on the tonal aptitude test. Mota suggests that this aptitude test does not give information about the musical achievement of the subjects. However, it is possible that the instruction was poor, resulting in the children not achieving to the levels that their aptitudes would allow.

Summary

Much of the research regarding singing achievement has examined method of instruction, technique, age and pitch discrimination. The results have been mixed. While it may seem logical to assume that those who sing accurately can discriminate pitch accurately and have a high level of tonal aptitude, findings do not always support this.

Rutkowski's SVDM has been validated and used by many researchers to examine extent of singing achievement. Rutkowski did find a small relationship between singing voice and aptitude. Atterbury and Silcox (1993) found that aptitude plays a role in singing ability, as did Conkling (1994). This supports Gordon's premise that the development of a child's singing achievement is determined by their aptitude.

Singing achievement is important to music educators. Researchers have investigated how and why students attain singing ability. Early studies used many varieties of aptitude tests and researcher-created singing tests. Only in the last ten years have the instruments begun to be standardized. With Gordon's aptitude tests and

Rutkowski's SVDM, researchers can begin to gather more uniform, valid, and easily interpretable information.

CHAPTER III

DESIGN AND METHODOLOGY

Participants

Fourth grade students (n=56) from three intact classes of a public elementary school in Michigan participated in this study. They were from a large, affluent, suburban area of Detroit. As part of the regular curriculum, the students receive music instruction twice a week for 30 minutes each session. The researcher is also the students' music teacher.

Since the 1997-1998 school year, the music curriculum for fourth grade students in the school has included Orff methodology, movement, and singing. From 1998 to the present, Gordon Music Learning Theory practices such as tonal and rhythm pattern instruction, and aural skill development have been added to the curriculum.

Design

This study is an examination of the relationship between a student's singing ability and tonal aptitude. Intermediate Measures of Music Audiation (IMMA) (Gordon, 1982) a standardized aptitude test, was used to measure the developmental, tonal aptitude of the students. IMMA was used for this study because it is designed for students in grades one through four and has also been used with some success with students in grades five and six. Primary Measures of Music Audiation (PMMA) (Gordon, 1979) was designed for students in kindergarten through grade three. Musical Aptitude Profile (MAP) (Gordon, 1995) was designed to test stabilized music aptitude of students in grades four through twelve; however MAP takes much longer to administer than IMMA, making it less practical for this study. The subjects took IMMA in the fall, making them young fourth

grade students. Fourth grade students were chosen for this study because fourth grade students' music aptitudes are at the end of the developmental stage and approaching stabilized music aptitude. It was decided that IMMA would be more appropriate than MAP or PMMA because of the age of the students and when they would take the test.

IMMA is designed to be administered to children in a group setting. It consists of two subtests, tonal and rhythm. The tonal subtest was used in this study to measure tonal aptitude. Because Rutkowski (1996) found a weak relationship between rhythm aptitude and singing achievement, the rhythm portion of IMMA was not administered to the students in this study.

When taking IMMA, children listen to a cassette tape and identify whether a tonal pattern is like or different from a second tonal pattern. The test has forty pairs of tonal patterns. The tonal patterns are performed in the same tempo but without rhythm. The examples contain both major and minor patterns. If the examples are the same, children draw a circle around two smiling faces on the answer sheet; if different, they draw a circle around two faces that were different, one smiling and one frowning.

The tonal subtest of IMMA takes approximately twenty minutes to administer. The verbal instructions are standardized. Gordon's aptitude tests have been subjected to extensive testing and research. IMMA tonal scores and their equivalent percentile rank scores are provided in the norms in the test manual.

Rutkowski's (1986) Singing Voice Development Measure (SVDM) was used to rate singing achievement (see Figure 5). The SVDM has been "shown to be a valid instrument to measure children's use of singing voice" (Rutkowski, 1990, p. 91). Research on singing ability suggests that between third and fourth grade, singing

achievement stabilizes (Cooper, 1995). Researchers report that older students sing more accurately than younger students (Yarbrough et.al, 1991), and vocal ranges of children have been shown to naturally increase with age (Feierabend; 1984, Jersild, 1934).

However, others such as Cooper (1995) disagree, “singing accuracy varies widely within and across grade levels” (p.229).

SINGING VOICE DEVELOPMENT MEASURE (SVDM)	
1.	Presingers: Children who do not sustain tones, but chant the text.
2.	Speaking-range singers: Children who sustain tones and exhibit some sensitivity to pitch, but remain within the speaking-voice range, usually A(2) to C(4) middle C.
3.	Uncertain singers: Children who sustain tones, but often waiver between a speaking voice and a singing voice range. When in singing voice, they use a range up to approximately F [#] (3) and seem to have difficulty lifting the voice above this pitch.
4.	Initial-range singers: Children who have the use of the singing voice range up to the register lift, usually A(3). At this stage, the children rarely drop back into speaking-voice range.
5.	Singers: Children who are able to sing over the register lift, B(3) and above, and have full use of their singing voices.

Figure 5. Rutkowski’s (1986) SVDM designed to assess children’s singing voice.

SVDM was developed for kindergarten students; it would be likely that many fourth grade students had already found their singing voice. The students in this study had experienced music once a week for thirty minutes up until the current year and did not have music instruction in kindergarten or before. This researcher believed that many of the subjects in this study had not found their singing voice from observation and knowledge of the students’ backgrounds.

A short, minor song used in Rutkowski's 1996 study was used as the criterion song (see Figure 5). Rutkowski (1990) stated that there has been some evidence that singing in a minor key is easier than singing in major key.



Figure 6. Criterion Song for Singing Achievement Assessment.

There have been several studies regarding the use of text or no text with regards to singing accuracy (Levinowitz, 1989, Goetze, 1985). There is some evidence that no text may produce greater singing accuracy, especially with children younger than third grade (Rutkoswki, 1987). However, many songs are sung with text and this study involves fourth grade students. For these reasons, it was decided that the subjects would sing the criterion song both with and without text.

Two weeks (four music class sessions) before the testing, students were taught the criterion song. A five to ten minute lesson was included in each of the fourth grade music classes to teach the song. The lessons focused on encouraging solo singing of the song so that students became comfortable singing alone. As suggested by Rutkowski (1990), the students were taped recorded while they sang in a group and had the recording played back for modeling of the eventual process and to ease anxiety about being tape-recorded.

Procedure

In November 1999, the researcher administered the tonal portion of IMMA, (Gordon, 1982) and tape recorded individual students singing the criterion song. It took the researcher two weeks to administer the singing achievement test to all students. Two

weeks prior to administering SVDM, the teacher sent letters to the parents inviting students to participate in the study. The letter included the nature of the testing and assurance of confidentiality of test results. At the time of the study, there were 68 students enrolled in the fourth grade. Fifty-six students participated in the study. Twelve students did not participate in the study; five did not return permission slips and six were not included in the study due to special education classification, (these students have a range of disabilities, such as, speech impairments, physical handicaps, and learning disabilities). Only one student's parents denied permission. This study received approval by the University Committee on Research Involving Human Subjects at Michigan State University, IRB#99429, Category: 2- G (see Appendix B).

The week before taking IMMA, the students were shown how the chairs would be arranged in the room and how they would be seated. The teacher explained that they would be given a clipboard, pencil, and paper. The students were told that they would listen to a tape recording and respond to what they hear. The teacher assured the students that names would be kept confidential and no one would know the results of their test. Also, the students were told that the test required no preparation on their part, and one class period would be used for test administration. No training for the test or further discussion concerning IMMA took place.

The criterion song for the singing achievement test was taught to the class for two weeks before the testing began. The fourth grade classes meet twice a week, so there were four instructional periods before the actual testing began. Students were taught the song in a rote, unison, and echo manner. The researcher used five to ten minutes at the beginning of each instructional period to echo-sing the song with the class. Students were

invited to sing voluntarily the song in solo into a toy microphone. In each of the four classes, students were randomly chosen and the researcher recorded the names of the students who performed so that eventually all of the students who participated in the study were given the opportunity to sing individually the criterion song before testing. Solo singing is encouraged in the regular curriculum, so this was not an anomaly. Solo singing was encouraged so that the students could hear themselves.

Teaching Procedure for the Criterion Song

The criterion song was taught using the following procedure:

Session 1

1. At the beginning of the 30-minute lesson period, the teacher asked the students to listen to the criterion song, which was sung without text on the syllable “bum.”
2. The teacher asked the students to echo-sing the first two measures of the song using the syllable “bum” in unison and then echo-sing the second two measures of the song.
3. The students then were asked to listen to the whole song, which was performed by the teacher using the text.
4. The students were asked to perform the entire song in unison using text.

Session 2

1. Repeat steps one through four from Session 1.
2. Several students were asked to sing the song individually. A toy microphone was used to make the activity more enjoyable.

3. The teacher showed the students the real microphone and tape recorder that would be used for testing. The teacher modeled how the students should sing into the microphone.
4. Students were assigned numbers, the teacher demonstrated how the student would speak the number and then perform the criterion song.

Session 3

1. Repeat steps one through four from Session 1.
2. Several students were asked to sing the song individually.

Session 4

1. Repeat steps one through four from Session 1.
2. The teacher showed the students the microphone and tape recorder that would be used for the testing. The teacher modeled how the students should sing into the microphone.
3. The students reviewed their assigned numbers. The teacher demonstrated how the student should speak their number and then perform the criterion song.
4. Several students were asked to sing the song individually.

The procedure used for the individual testing of singing achievement was explained to students on the fourth class session.

The testing of the singing achievement began the same week that the students took IMMA. Testing of students was agreed upon and arranged between the researcher and the classroom teachers. The students from the three fourth grade classrooms were mixed randomly so that individual classes were not intact. Students were sent two at a time to

the music room, a familiar room, where they were tape-recorded. They were instructed to wait outside until the student ahead of them left. Each student was assigned a number to insure confidentiality. The researcher played the first three notes of the song and then said “ready sing.” The students sang the song individually, once with, and once without the text and the researcher tape-recorded the performances. The song was performed first without text followed by with text.

Two judges were asked to listen to the tape and independently rate the singing performances using SVDM. The judges were chosen based on their vocal study and familiarity with the elementary student voice. Both judges are vocal music education teachers in public school districts in Michigan. One judge is currently teaching elementary general music, the other taught elementary general music until 1999 and is currently a middle school choral director.

The researcher contacted each judge by phone and described the judging procedure. Following this conversation, a packet was sent to the judges containing the following: a duplicate copy of the tape of the singing achievement test performances, a judge’s scoring sheet for the SVDM (see Appendix A), a copy of the criterion song, and rating sheets for each student. Also included was a letter describing the numbering sequence of the students and instructions for giving two scores to each student, one for singing without text, and one for singing with text. The judges were asked to study SVDM before rating the students. They were asked to listen carefully to each student, stop the tape, and give a score from the continuous rating scale of one to five (Appendix A). The tape itself was thirty minutes long, and it was estimated that the entire process

would take each judge from thirty minutes to an hour. The researcher recommended that the judges complete the scoring at one time and not over an extended period.

Design and Analysis

The means, standard deviations, and standard errors of measurement were calculated for the tonal aptitude scores and the singing accuracy scores. Inter-judge reliability was calculated and split halves reliabilities were calculated for the tonal aptitude scores. Pearson r was used to find the relationship between tonal IMMA and SVDM scores.

CHAPTER IV

ANALYSIS AND INTERPRETATION OF DATA

IMMA was administered to measure the tonal aptitude of the students, and SVDM was administered to measure their level of singing achievement. The results are presented as follows: (1) means and standard deviations and the standard errors of measurement for tonal aptitude scores and singing accuracy scores; (2) split-halves reliabilities for the tonal aptitude scores; (3) interjudge reliability of the judges ratings for singing accuracy; and (4) Pearson r for the relationship between the tonal subtest of IMMA and SVDM scores.

Results

Means, Standard Deviations, and Standard Errors of Measurement for IMMA and SVDM

A comparison of the means, standard deviations, and standard errors of measurement of the sample population with those reported in the IMMA manual show that the scores of the sample in this study function similarly to those of the standardization sample (see Table 1).

A comparison of the means and standard deviations show the mean to be identical and standard deviations to be only slightly higher for the standardization sample. Standard error of measurement, however, is smaller in the study sample than in the standardization sample.

Table 1. Means, Standard Deviations and Standard Errors of Measurement for the Tonal Subtest of IMMA.

	M	SD	Std. Error
Study sample	35.2	2.9	.4
Standardization sample	35.2	3.0	1.1

The means, standard deviations, and standard errors of measurement for the singing scores are given as singing with text, singing without text, and a composite singing score (see Table 2). These means, standard deviations and standard error of measurement represent a combination of the two judges scores. These means are higher than the theoretical means of six and twelve, but not dramatically so. The observed standard deviations are also higher than the theoretical standard deviations of 1.34 and 2.68.

Table 2. Means, Standard Deviations and Standard Errors of Measurement for Singing Voice Development Measure.

	M	SD	Std.Error
Singing with text	6.3	2.4	.3
Singing without text	6.6	2.2	.3
Singing Composite	12.9	4.5	.6

There is a small difference between the mean of the singing score with text and the singing score without text. Students who sang without text performed with slightly greater vocal accuracy than those who sang with text. This is particularly remarkable considering that all students performed without text first, so that the performance with text was with more practice.

Split-Halves Reliabilities of Tonal Subtest of IMMA

Split-halves reliabilities using Spearman Brown Formula was .89 for the tonal subtest of IMMA. That score compares favorably to what is reported in the manual, which was .72 with fourth grade students.

Interjudge Reliability of Singing With Text and Singing Without Text

Interjudge reliability between the two independent judges was determined for singing with text and singing without text. The interjudge reliabilities can be found in Table 3. Although these interjudge reliabilities are not high, they are acceptable.

Table 3. Interjudge Reliability for Singing Accuracy Scores With and Without Text and for Singing Composite Scores.

No Text	Text	Composite
.70	.73	.75

Correlations Between Singing Achievement and Tonal Aptitude

The relationship between student scores on singing achievement and tonal aptitude scores are shown in Table 4. The relationship between tonal aptitude scores and scores on singing without text is significant at the .05 level. The relationship between tonal aptitude scores and singing with text is significant at the .01 level. The relationship between tonal aptitude scores and composite singing scores is significant at the .01 level (see Table 4).

Table 4. Correlations between Scores for Singing Without Text, With Text, and Singing Composite, and Scores for Tonal Aptitude

No text	Text	Composite
.35*	.42**	.40**

* $p < .05$

** $p < .01$

Interpretations

This researcher expected that students who have a high tonal aptitude would have achieved a high level of singing accuracy by the nature of their ability.

Even though scores for singing without text and singing with text both were found to be significantly correlated with tonal aptitude, there was a difference between the strength of the relationship when singing without text and singing with text. This was possibly due to the method in which students sang the criterion song. All of the subjects were asked to sing the criterion song without words first and then with words. In retrospect, these two should have been alternated so that half of the students sang the criterion song without words first then with words and the other half of the students sang with words first then without the words.

A significant relationship between the singing accuracy scores and the tonal subtest of IMMA may be the result of several factors. The quality of instruction may determine the success with which a student engages SVDM. SVDM measures a student's ability to sing in their head voice as opposed to chest voice. If a teacher does not give instruction in using head voice then students cannot achieve, regardless of their level of music aptitude.

CHAPTER V

SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS

Summary

This study was conducted to determine if a relationship exists between singing achievement and tonal aptitude.

A review of related literature revealed that several studies have been conducted that are similar to this one. These studies used Rutkowski's SVDM to measure singing voice development and the tonal section of Gordon's PMMA to measure tonal aptitude.

Rutkowski's (1996) year-long study revealed a small relationship between singing voice and developmental tonal aptitude. A one year study by Atterbury and Silcox (1993) found that students with high aptitude also had a high level of singing ability. Conkling (1994) also found that students who have a high aptitude had higher levels of achievement than those students who have a moderate or low aptitude. Mota (1997) found no relationship between PMMA and music tasks, which included two singing tasks.

Fifty-six fourth grade students from a public school in Michigan participated in this study. The students in the study receive music instruction twice a week for thirty minutes each session.

The students were tested using the tonal section of Intermediate Measures of Music Audiation (IMMA). The performances of the criterion song, which was taught two weeks before, were tape recorded and evaluated by two judges using SVDM. Interjudge reliability between the judges established that there was consistency of the singing evaluation between the two judges. The relationship between singing achievement and tonal aptitude was statistically significant at the .01 level.

Conclusions

All but one of the studies that examined the relationship between singing achievement and tonal aptitude indicated that there is a small relationship between the two variables. Mota (1997) did not find a relationship between aptitude and achievement tasks. Rutkowski (1996), Atterbury and Silcox (1993), and Mota (1997) used Gordon's PMMA and measured developmental aptitude with kindergarten students. Conkling (1994) used Gordon's Advanced Measures of Music Audiation to measure stabilized aptitude with seventh grade students.

This study measured the stabilized music aptitude of fourth grade students using Gordon's Intermediate Measures of Music Audiation. The results indicate that students who have high tonal aptitude also tend to be able to use their singing voices as measured by SVDM. This would support the statement by Gordon that "the degree to which good tonal habits are developed, however, is dependent upon tonal aptitude" (1971, p. 93).

No other duplication of this study that the researcher is aware of exists. Several studies exist that are similar to this study. The study by Mota is perhaps the closest, but Mota's three-year study examined aptitude as measured by PMMA and various musical tasks, including singing. She did not use SVDM and confused the aptitude results with achievement results, stating that the aptitude test must have been incorrect because it did not measure the achievement of the students. It is also possible that the students had not received appropriate instruction, so their aptitudes were not reflected in their achievement.

Rutkowski's results after administration of the tonal subtest of PMMA, singing instruction with treatment groups in small and individual group settings, and a control

group of singing instruction with large group settings revealed a small relationship between developmental music aptitude and singing achievement. The results of the PMMA did not change over a pre and posttest period but, “both groups had gains in tonal aptitude mean scores during the treatment” (Rutkowski, 1996, p.361). She concluded that the correlation coefficients were very low from SVDM and PMMA, with posttest scores of .20 indicating a “very small relationship between use of singing voice and developmental tonal aptitude” (1996, p. 361). This study found a relationship between singing voice and stabilized tonal aptitude, which could account for the difference between this study and Rutkowski’s.

Recommendations

The following recommendations for further research are based on this research and are concerned with singing achievement and the use of aptitude scores.

SVDM has been validated for use with kindergarten students. However, the rating scale may be inadequate for use with fourth grade students. Further researchers may wish to expand the range of this measurement into intonation accuracy.

A standardized measurement for singing accuracy with older children is not available. One should be developed. Then, it would warrant further research with regards to the relationship between the singing accuracy and tonal aptitude of students who have stabilized music aptitude.

This study could be replicated, fixing the design flaw related to text and no text. The students were asked to sing the no text example first then to sing the text. Alternating the performance order of text and no text would control for the order effect in the testing procedure.

The judge's training for this study was done by phone and written correspondence. While interjudge reliability was acceptable, it is suggested that if replicated the judges receive training with the researcher in person.

It is recommended that replication of this study involve a larger and perhaps more heterogeneous population. Additional studies should also consider the difference between schools where students have music instruction twice a week from kindergarten and beyond and schools that have music instruction once a week without kindergarten instruction.

Implications

A relationship between tonal aptitude as measured by IMMAT and singing achievement as measured by SVDM could also indicate that IMMA is a valid measure of students' tonal aptitude scores. If tonal aptitude is predictive of tonal achievement, it naturally follows that a valid measure of tonal aptitude should relate to scores on a valid measure of tonal singing achievement.

Singing is a learned behavior. The results of this study suggest that singing achievement is related to tonal aptitude. This information can help teachers assess the strengths and weaknesses of their students. However, because the relationship is not perfect, teachers should not assume that a child who cannot sing has a low tonal aptitude. Students with a high tonal aptitude may have low singing achievement because they have not learned how to sing. The extent to which a teacher provides singing instruction may determine the success a student can achieve. If a teacher doesn't give singing instruction, students cannot be expected to achieve. That a relationship exists may suggest that singing supports developmental aptitude and vice versa.

APPENDIX A

Singing Voice Development Measure

Student Number _____ (please verify the student on the tape recording).

SINGING VOICE DEVELOPMENT MEASURE (SVDM)

1. Presingers: Children who do not sustain tones, but chant the text.
2. Speaking-range singers: Children who sustain tones and exhibit some sensitivity to pitch, but remain within the speaking-voice range, usually A(2) to C(4) middle C.
3. Uncertain singers: Children who sustain tones, but often waiver between a speaking voice and a singing voice range. When in singing voice, they use a range up to approximately F[#](3) and seem to have difficulty lifting the voice above this pitch.
4. Initial-range singers: Children who have the use of the singing voice range up to the register lift, usually A(3). At this stage, the children rarely drop back into speaking-voice range.
5. Singers: Children who are able to sing over the register lift, B_♭(3) and above, and have full use of their singing voices.

By Joanne Rutkowski

Please rank the student with a number from 1-5 _____

Judge Number _____

APPENDIX B

Letter to parents for permission to participate in the study

Dear Parents and Guardians,

I am currently a graduate student at Michigan State University. As part of my course work, I will be administering a music aptitude profile to students in the fourth grade. Following the aptitude evaluation I will listen to each student, individually, and administer a singing achievement test. Both tests are voluntary. The results will not contain students' names. Using information from the profile and achievement test I will be better able to address the individual needs of each student.

Please complete the form at the bottom of this page. If you have any questions regarding this project please call me at Eagle Elementary or my voice mail at 426-1190. Any further questions you may have regarding being a subject in this study, please contact, David E. Wright, UCRIHS Chair, 1-517-355-2180. Thank you.

Sheri Jaffurs, Music Teacher
Eagle Elementary School

*****Detach*****

_____ has permission to take the Music
Name of Student

Aptitude Profile and the Singing Achievement Test.

Signature of Parent or Guardian

LIST OF REFERENCES

Apfelstadt, H. (1984). Effects of melodic perception instruction on pitch discrimination and vocal accuracy of kindergarten children. Journal of Research in Music Education, 32(1), 15-24.

Apfelstadt, H. (1988). What makes children sing well? Update: The application of research in music education, 7(1-2), 27-32.

Atterbury, B. W., & Silcox, L. (1993). The effect of piano accompaniment on kindergartners' developmental singing ability. Journal of Research in Music Education, 41(1), 40-47.

Ball, W. (1995). Nurturing musical aptitude in children. Dimensions of Early Childhood, 23(4), 19-21, 23-24.

Boardman, E. (1964). An investigation of the effect of preschool training on the development of vocal accuracy in young children. (Doctoral dissertation, University of Illinois, Urbana-Champaign, 1964). Dissertation Abstracts International, 0090, 25-02.

Carlsen, J. (1981). Applications of psychology to the teaching and learning of music. Documentary Report of the Ann Arbor Symposium. Reston, Virginia: MENC.

Conkling, S. (1994). A comparison of the effects of learning sequence activities and vocal development exercises on the vocal music achievement of middle level students (Doctoral dissertation, University of Rochester, Eastman School of Music, 1994). Dissertation Abstracts International, UMI No. DA9503122.

Cooper, N. A. (1995). Children's singing accuracy as a function of grade level, gender, and individual versus unison singing. Journal of Research in Music Education, 43(3), 222-231.

- Deutsch, D. (Ed.). (1982). The psychology of music. New York: Academic Press.
- Dunn, L. (1959). Peabody picture vocabulary test. Minnesota: American Guidance Service, Inc.
- Feierabend, J. (1984). The effects of specific tonal pattern training on singing and aural discrimination abilities of first grade children. (Doctoral dissertation, Temple University, Philadelphia, 1984) University Microfilms Order No. 6977335.
- Feierabend, J. (1990, July, August). Music in early childhood. Design for Arts in Education, 91(6), 15-19.
- Gardner, H. (1983). Frames of mind. (2nd ed.). New York: Basic Books.
- Gardner, H. (1999). Intelligence reframed. New York: Basic Books.
- Goetze, M. (1985). Factors affecting accuracy in children's singing. (Doctoral dissertation, University of Colorado at Boulder, 1985). University Microfilms Order No. 8528488.
- Goetze, M., & Horii, Y. (1989). A comparison of the pitch accuracy of group and individual singing in young children. Bulletin of the Council for Research in Music Education, 99, 57-73.
- Goetze, M., Cooper, N., & Brown, C. M., (1990). Recent research on singing in the general music classroom. Bulletin of the Council for Research in Music Education, 104, 16-37.
- Gordon, E. E. (1971). The psychology of music teaching. Englewood Cliffs, New Jersey: Prentice-Hall.
- Gordon, E. E. (1982). Intermediate measures of music audiation. Chicago: G.I.A. Publications.

Gordon, E. E. (1986, 1979). Manual for the primary measures of music audiation and the intermediate measures of music audiation. Chicago: G.I.A. Publications.

Gordon, E.E. (1997). Learning sequences in music. (5th ed.). Chicago: GIA. Publications.

Gordon, E. E. (1997). A music learning theory for newborn and young children. Chicago: GIA. Publications.

Gould, A.O. ((1969). Developing specialized programs for singing in the elementary schools. Bulletin of the Council for Research in Music Education, 17, 9-22.

Green, G. A. (1989). The effect of vocal modeling on pitch-matching accuracy of elementary school children. Journal of Research in Music Education, 38(3), 225-231.

Greenburg, M. (1979). Your children need music. Englewood Cliffs, New Jersey: Prentice-Hall.

Hutchins, P. (1972). Good night, owl! New York: Macmillan Publishing Company, Inc.

Jersild, A., & Bienstock, S. (1934). A study of the development of children's ability to sing. The Journal of Educational Psychology, 25 (7), 481-503.

Jones, W. M. (1994). Carol Elaine Rupp: the effects of vocal modeling and melodic direction on development of head voice placement in 4-year old nonsinging children. Bulletin of the Council for Research in Music Education, 120, 84-86.

Joyner, D. R. (1969). The monotone problem. Journal of Research in Music Education, 17(1), 115-124.

Kramer, S. (1985). The effects of two different music programs on third and fourth grade children's ability to match pitches vocally. (Doctoral dissertation, Rutgers, The State University of New Jersey, 1985).

Lange, D. (1999). The effect of the use of text in music instruction on the tonal aptitude, tonal accuracy, and tonal understanding of kindergarten students. (Doctoral dissertation, Michigan State University, 1999).

Levinowitz, L. M. (1987). An experimental study of the comparative effects of singing songs with words and without words on children in kindergarten and first grade. (Doctoral dissertation, Temple University, 1987). Dissertation Abstracts International. 48, 65.

Levinowitz, L. M. (1989). An investigation of preschool children's comparative capability to sing songs with and without words. Bulletin of the Council for Research in Music Education, 100, 14-19.

Levinowitz, L. M., Barnes, P., Buerrini, S., Clement, M., D'April, P., & Morey, M. J. (1998). Measuring singing voice development in the elementary general music classroom. Journal of Research in Music Education, 46(1), 35-47.

Mota, G. (1997, Summer). Detecting young children's musical aptitude: A comparison between standardized measures of music aptitude and ecologically valid musical performances. Bulletin of the Council for Research in Music Education, 133, 89-94.

Petzold, R. G. (1963, Spring). The development of auditory perception of musical sounds by children in the first six grades. Journal of Research in Music Education, 11, 21-43.

Phillips, K.H. (1983). The effects of group breath control training on selected vocal measures related to the singing ability of elementary students in grades two, three and four. (Doctoral dissertation, Kent State University, 1983). Dissertation Abstracts International, 44A, 1017.

Phillips, K. H. (1989). Mary Goetze: Factors affecting accuracy in children's singing. Bulletin of the Council for Research in Music Education, 102, 82-85.

Phillips, K. H. (1992). Teaching kids to sing. New York: Schirmer.

Phillips, K. H. (1993). Back to basics: Teaching children to sing. General Music Today, 6(3), 30-32.

Phillips, K. H., & Aitchison, R. (1997). Effects of psychomotor instruction on elementary general music students' singing performance. Journal of Research in Music Education, 45(2), 185-197.

Radocy, R., & Boyle J. D. (1997). Psychological foundations of musical behavior (3rd ed.). Springfield, Illinois: Charles C. Thomas Ltd.

Ramsey, J. H. (1983). The effects of age, singing ability, and instrumental experiences on preschool children's melodic perception. Journal of Research in Music Education, 31(2), 133-145.

Rupp, C. E. (1994). The effects of vocal modeling and melodic direction on development of head voice placement in 4-year-old nonsinging children (Doctoral dissertation, University of Missouri, 1992). Dissertation Abstracts International, 53, 04A.

Rutkoswki, J. (1985). The child voice: An historical perspective. The Bulletin: Historical Research in Music Education, 6(1), 1-15.

Rutkowski, J. (1986). The effect of restricted song range on kindergarten children's use of singing voice and developmental music aptitude. (Doctoral dissertation, State University of New York at Buffalo, 1986).

Rutkowski, J. (1989). The problem singer: What does research suggest? General Music Today: The Journal of the Society for General Music, 2(3), 24-29.

Rutkowski, J. (1990). The measurement and evaluation of children's singing voice development. The Quarterly Journal of Music Teaching and Learning, 1(1-2), 81-95.

Rutkowski, J. (1996). The effectiveness of individual/small-group singing voice and developmental music aptitude. Journal of Research in Music Education, 44(4), 353-368.

Swears, L. (1985). Teaching the elementary school chorus. West Nyack, New York: Parker.

Webster's II new riverside dictionary (3rd ed.). (1984). New York: Berkley Books.

Welch, G. (1979). Poor pitch singing: A review of the literature. Psychology of Music, 7(1), 50-58.

Yarbrough, C., Green, G., Benson, W., & Bowers, J. (1991). Innacurate singers: An exploratory-search of variables affecting pitch-matching. Bulletin of the Council for Research in Music Education, 107, 23-34.

Youngson, S. C., & Persellin, D. C. (1999, November). The kinesthetic effect of Curwen hand signs on vocal accuracy of young children. Poster session presented at the American Orff Schulwerk Association National Conference, Phoenix, AZ.

MICHIGAN STATE UNIV. LIBRARIES



31293020799270