THE EFFECT OF SCHOOL PERFORMING ENSEMBLE PARTICIPATION ON THE ABILITY TO PERFORM AND PERCEIVE EXPRESSION IN MUSIC

DISSERTATION

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

By

Daryl Wayne Kinney, B.M., M.M.

*****

The Ohio State University
2001

Dissertation Committee:

Dr. Jere L. Forsythe, Advisor
Dr. Patricia J. Flowers
Dr. David B. Huron
Dr. Russel C. Mikkelson

Approved by

Adviser
School of Music
Copyright by
Daryl Wayne Kinney
2001
ABSTRACT

The purpose of this study was to examine the effects of school performing ensemble experience on the ability to perceive expression in music and to perform a musical passage expressively. The study compared collegiate level subjects who had prior experience in high school performing ensemble courses with subjects who had not. The study was divided into two phases: Phase I addressed the performance of musical expression; Phase II examined the perception of musical expression. Because expression is contingent to a large degree on the accuracy of musical performance, an examination of musical accuracy was included in each phase as well.

In Phase I, an experiment was devised in which subjects performed two familiar songs and one unfamiliar song using a MIDI-based, minimal skill musical device under two expressive conditions: expressively and unexpressively. Subjects were divided for analysis into musicians (N=28) and nonmusicians (N=28) based upon their musical background. Musicians were subjects whose only musical experience was in school performing ensembles; nonmusicians were subjects who had no prior training in music. The 336 performances generated (56 subjects X 6 performances each) were divided among four expert musicians for adjudication. Judges rated performances for musical expression and accuracy.
In Phase II, 54 performances from Phase I were randomly selected for use in an experiment devised to examine the perception of musical expression and accuracy. Subjects, musicians (N=73) and nonmusicians (N=56), listened to the performances and rated each for musical expression and accuracy.

Results indicated that musicians showed significant differences between their expressive and unexpressive performances, and were able to perform more accurately than nonmusicians. Although nonmusicians were able to demonstrate a difference between expressive and unexpressive performances, this difference was not significant. Furthermore, musicians' expressive performances were judged to be significantly more expressive than those of nonmusicians. In regard to perception, musicians were more discriminating and more consistent in their ratings than nonmusicians, and were found to rate performances similar to those of the judges from Phase I. Musicians demonstrated significant differences between their ratings of expressive and unexpressive performances. No significant difference was evidenced between nonmusicians ratings of expressive and unexpressive performances.
To Kimberly
ACKNOWLEDGMENTS

As with all things in life, this project is the result of the efforts of many. In no way can the small amount of words contained on these pages begin to repay the debt that I owe each individual for their enduring love, support, inspiration, faith, and prayers.

I would first like to give thanks to God who makes all things possible.

To my beautiful wife Kimberly who has read every word I have ever written (and has stayed with me in spite of it), this and all else in my life is possible because of you. Thank you for your love, patience, understanding, encouragement, and faith.

I am eminently grateful to my wonderful family who has supported Kimberly and me throughout this entire endeavor. A special thank you is owed to my father who provided the monetary resources to make much of this project possible. I am humbled and truly blessed to have you in my life. Thank you for always believing in your son.

To my adviser, mentor, teacher, and friend, Jere Forsythe, I cannot begin to express the profound impact that you have had on my life and thinking as a teacher, musician, scholar, researcher, and member of the human race. I am forever in your debt. Thank you for three years of the most stimulating, intellectual, and enjoyable experiences of my life. I truly am privileged to have been your student.
I am grateful and indebted to Pat Flowers who has been much like a second adviser for me. The impact that your knowledge, passion for research, and belief in the scholarly enterprise has had on me is tremendous. Thank you for always being there and for making house calls!

To Russ Mikkelson, my heartfelt and sincerest thanks. Thank you for your inspiration, insight, and belief. The opportunities that you provided for me have helped me grow and develop in all areas of my teaching. I am a better musician and conductor because of you. Our time together has meant more than you can imagine.

I would also like to thank David Huron for inspiring me to pursue my interests in music perception and cognition. Without your dynamic and thought provoking teaching, this path may never have been considered. Thank you for your advice and encouragement along the way.

To John Kennedy, my statistical hero, thank you for making statistics fun and exciting. I am honored to have been your student. This dissertation never would have been possible without the skills you taught me.

A special thank you is owed to those who helped in the logistics of the project. First, thank you to Dr. Daniel Avorgbedor and his teaching assistants, Josh Veltman, Eric Berg, Tom Beardslee, and Kristen Holland, who allowed me to use valuable class time in their “Music Cultures of the World” courses in order to recruit subjects and administer the experiment. Knowing how valuable teaching time is, I cannot thank you enough. Second, to those who proctored the experiment, Dr. Jan McCrary, Dr. Jacqueline Henninger, Marla Butke, and Carlos Abril, thank you for being so giving of your time. I would have no data (and, hence, no dissertation) if it weren’t for you! Third, I am grateful to my expert evaluators, Dr. Mark Ellis, Dr. Sandra Guerard, Dr. David Tovey, and Dr. Paul
Sanders. Thank you for taking the time out of your busy schedules to help in the project. And lastly, thank you to Ray Spillman and Pete Tender who not only found a room that could be used specifically for the experiment, but also helped set up all of the MIDI stations and helped trouble shoot any problems I encountered along the way. I could not have gotten the project off the ground without your assistance.

Finally, I would like to thank the Music Education faculty at the Ohio State University. The leadership and inspiration that I received from all of you will remain with me for the rest of my life. Thank you......always.
VITA

March 20, 1971................................. Born- Richmond, Virginia

1994............................................. B.M. (Music Education)
Magna Cum Laude
Virginia Commonwealth University

1994-1998...................................... Director of Bands
Lloyd C. Bird High School
Salem Middle School
Chesterfield, Virginia

1998............................................. M.M. (Music Education)
Virginia Commonwealth University

1998-2001...................................... Graduate Teaching Associate
The Ohio State University
Columbus, Ohio

FIELDS OF STUDY

Major Field: Music Education

Studies in Music Education: Professors Jere L. Forsythe,
Patricia J. Flowers, Timothy Gerber, Robert Gillespie,
Jan McCrary, Patricia O' Toole, and Jon Woods

Studies in Conducting: Professors Russel Mikkelsen and
Richard Blatti
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abstract</td>
<td>ii</td>
</tr>
<tr>
<td>Dedication</td>
<td>iv</td>
</tr>
<tr>
<td>Acknowledgments</td>
<td>v</td>
</tr>
<tr>
<td>Vita</td>
<td>viii</td>
</tr>
<tr>
<td>List of Tables</td>
<td>xi</td>
</tr>
<tr>
<td>List of Figures</td>
<td>xiii</td>
</tr>
<tr>
<td><strong>Chapters:</strong></td>
<td></td>
</tr>
<tr>
<td>1. Introduction</td>
<td>1</td>
</tr>
<tr>
<td>Background</td>
<td>1</td>
</tr>
<tr>
<td>Need</td>
<td>7</td>
</tr>
<tr>
<td>Statement of the Problem</td>
<td>10</td>
</tr>
<tr>
<td>Purpose</td>
<td>11</td>
</tr>
<tr>
<td>Definition of Terms</td>
<td>12</td>
</tr>
<tr>
<td>Limitations</td>
<td>15</td>
</tr>
<tr>
<td>Overview of Remaining Chapters</td>
<td>15</td>
</tr>
<tr>
<td>2. Review of Literature</td>
<td>17</td>
</tr>
<tr>
<td>Studies in the Performance of Musical Expression</td>
<td>21</td>
</tr>
<tr>
<td>Studies in the Perception of Musical Expression</td>
<td>21</td>
</tr>
<tr>
<td>Research into the Effect of Instruction on the Perception and Production of Musical Nuance</td>
<td>33</td>
</tr>
<tr>
<td>Summary</td>
<td>53</td>
</tr>
</tbody>
</table>
3. Methods and Procedures
   Purpose..............................................................................................................55
   Participants......................................................................................................57
   Demographics..................................................................................................57
   The Experimental Apparatus..........................................................................59
   Selection of the Musical Excerpts.................................................................59
   Method: Phase I...............................................................................................61
   Method: Phase II.............................................................................................67
   Analysis of Data...............................................................................................69

4. Results............................................................................................................71
   The Effect of Order, Gender, and Type of Ensemble Participation..................73
   Inter-rater Reliability and Internal Consistency..............................................74
   Presentation of the Data from Phase I............................................................77
      Research Question One..............................................................................77
      Research Question Two...............................................................................84
      Research Question Three...........................................................................89
      Research Question Four...............................................................................95
   Presentation of the Data from Phase II..........................................................100
      Research Question Five.............................................................................100
      Research Question Six..............................................................................106
      Research Question Seven..........................................................................111
      Research Question Eight............................................................................116

5. Summary, Discussion, Conclusions, and Recommendations for Further Research........................................................................................................121
   Summary.........................................................................................................121
   Discussion.......................................................................................................128
   Conclusions....................................................................................................139
   Recommendations for Further Research....................................................141

Bibliography........................................................................................................143

Appendices.........................................................................................................156
   Appendix A - Musical Excerpts.................................................................156
   Appendix B - Musical Background Survey..................................................160
   Appendix C - Post Experiment Follow Up Survey.......................................164
   Appendix D - Likert Scale Response Form..................................................166
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Means (M) and Standard Deviations (SD) of Expression Scores for Musician and Nonmusician Performers Reported by Song and Expressive Condition</td>
<td>78</td>
</tr>
<tr>
<td>2. Means (M) and Standard Deviations (SD) of Expression Scores for Musician and Nonmusician Performers Collapsed Across Songs</td>
<td>79</td>
</tr>
<tr>
<td>3. 2 x 3 x 2 Repeated Measures Analysis of Variance of Expression Scores by Group (Musicians vs. Nonmusicians)</td>
<td>81</td>
</tr>
<tr>
<td>4. Means (M) and Standard Deviations (SD) of Accuracy Scores for Musician and Nonmusician Performers Reported by Song and Expressive Condition</td>
<td>85</td>
</tr>
<tr>
<td>5. 2 x 3 x 2 Repeated Measures Analysis of Variance of Accuracy Scores by Group (Musicians vs. Nonmusicians)</td>
<td>87</td>
</tr>
<tr>
<td>6. Means (M) and Standard Deviations (SD) of Expression Scores for Number of Years in a School Performing Ensemble as Reported by Song and Expressive Condition</td>
<td>90</td>
</tr>
<tr>
<td>7. Means (M) and Standard Deviations (SD) of Expression Scores by Number of Years of School Ensemble Participation Collapsed Across Songs</td>
<td>92</td>
</tr>
<tr>
<td>8. 6 x 3 x 2 Repeated Measures Analysis of Variance of Expression Scores by Number of Years of Participation in a School Performing Ensemble (Y)</td>
<td>94</td>
</tr>
<tr>
<td>9. Means (M) and Standard Deviations (SD) of Accuracy Scores by Number of Years Participation in a School Performing Ensembles Reported by Song and Expressive Condition</td>
<td>96</td>
</tr>
</tbody>
</table>
10. 6 x 3 x 2 Repeated Measures Analysis of Variance of Accuracy Scores by Number of Years of Participation in a School Performing Ensemble (Y) ................................................................. 98

11. Means (M) and Standard Deviations (SD) of Musicians’ and Nonmusicians’ Ratings of Perceived Expression as Reported by Song and Expressive Condition ......................................................... 101

12. 2 x 3 x 2 Repeated Measures Analysis of Variance of Ratings of Perceived Expression by Group (Musicians vs. Nonmusicians) .... 103

13. Means (M) and Standard Deviations (SD) of Musicians’ and Nonmusicians’ Ratings of Perceived Accuracy as Reported by Song and Expressive Condition ......................................................... 107

14. 2 x 3 x 2 Repeated Measures Analysis of Variance of Ratings of Perceived Accuracy by Group (Musicians vs. Nonmusicians) .... 109

15. Means (M) and Standard Deviations (SD) of Ratings of Perceived Expression by Numbers of Years Participation in a School Performing Ensemble as Reported by Song and Expressive Condition ...................................................................................... 112

16. 9 x 3 x 2 Repeated Measures Analysis of Variance of Ratings of Perceived Expression by Number of Years in a School Performing Ensemble ...................................................................................... 114

17. Means (M) and Standard Deviations (SD) of Ratings of Perceived Accuracy by Numbers of Years Participation in a School Performing Ensemble as Reported by Song and Expressive Condition ...................................................................................... 117

18. 9 x 3 x 2 Repeated Measures Analysis of Variance of Ratings of Perceived Accuracy by Number of Years in a School Performing Ensemble ...................................................................................... 119
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Mean Expression Ratings for Musicians and Nonmusicians Collapsed Across Songs</td>
<td>82</td>
</tr>
<tr>
<td>2.</td>
<td>Mean Expression Ratings for Songs Collapsed Across Group and Expressive Condition</td>
<td>83</td>
</tr>
<tr>
<td>3.</td>
<td>Mean Accuracy Ratings for Musicians and Nonmusicians Collapsed Across Expressive Condition by Song Performed</td>
<td>88</td>
</tr>
<tr>
<td>4.</td>
<td>Mean Expression Ratings by Number of Years of Participation in a School Performing Ensemble</td>
<td>93</td>
</tr>
<tr>
<td>5.</td>
<td>Mean Accuracy Ratings by Number of Years of Participation in a School Performing Ensemble</td>
<td>99</td>
</tr>
<tr>
<td>6.</td>
<td>Mean Rating of Perceived Expression Depicting Second Order Interaction Between Group, Song, and Expressive Condition</td>
<td>105</td>
</tr>
<tr>
<td>7.</td>
<td>Mean Ratings of Perceived Accuracy Depicting Interaction Between Song and Expressive Condition</td>
<td>110</td>
</tr>
<tr>
<td>8.</td>
<td>Mean Ratings of Perceived Expression Across All Years of School Ensemble Participation Depicting Interaction Between Song and Expressive Condition</td>
<td>115</td>
</tr>
<tr>
<td>9.</td>
<td>Mean Ratings of Perceived Accuracy Across All Years of School Ensemble Participation Depicting Interaction Between Song and Expressive Condition</td>
<td>120</td>
</tr>
</tbody>
</table>
CHAPTER 1

INTRODUCTION

Background

The performing ensemble constitutes one of the most seminal educational mediums in American music education today. From the most basic elementary school choirs to the most sophisticated collegiate full-orchestras, performing groups can be found at all levels of education. The current status of the performing ensemble should come as no surprise since it has enjoyed a rich and extensive history in American education and culture. Indeed, the performing ensemble, itself, has been considered by many to be synonymous with music education for most of the 20th and 21st centuries (Reimer, 1970).

The dominance of performing is nowhere more evident than in the secondary school. It is at this level where performing ensemble electives oftentimes monopolize course offerings. Although only 15%-25% of the school population usually elects to participate in these courses, these programs, and their musical successes, are regarded as an indication of what "can be achieved musically in an open education system" (Reimer, 1989, p. 182). That is, the high levels of excellence achieved in musical performance by school ensembles are interpreted as an indicator of preeminent musical
training taking place in American education. However, the assumption that
excellence in musical performance will somehow transfer directly to overall
musical learning and understanding has come under great scrutiny.

All too often the development of technical skills to perform at high
levels of musicianship is equated with and assumed to transfer to deeper
musical knowledge and understanding. In fact, skill development is
interpreted frequently “as clear evidence that [students’] understanding, their
sensitivity, their preferences, their attitudes are all progressing in the
direction of a deeper grasp of music” (Reimer, 1970, p. 130). Unfortunately,
research tends to indicate that there is little if any relationship between school
performance activities and, for example, subsequent tastes or preferences for
music. Mercer (1972) illustrated that students with four years of band
instruction were only slightly better prepared to listen to a symphony concert
than those who received no music instruction at all in high school.
Similarly, Duerksen (1968) reported that participation in performance
activities seemed to have only minor positive effects on listening skills.
These investigations led Benner (1972) to conclude that “performing group
participation has little effect on musical behavior other than the acquisition
of performance skills, unless there is planned effort by the teacher to enrich
the performing experience with additional kinds of musical understanding”
(p. 10).

The advent of the Comprehensive Musicianship (CM) concept at the
Northwestern Seminar in 1965 sought to redirect the emphasis of school
music performing courses and alleviate these programs’ shortcomings.
Developed under the auspices of the Contemporary Music Project (CMP), CM
initially was directed at undergraduate music programs in an effort to unify
and relate various aspects of the curriculum usually studied separately (e.g., theory, history, performance, composition, etc.). However, the greatest impact of the concept was eventually realized in elementary and secondary school performing programs where CM was implemented to “counter the Yale Seminar criticism that performance-oriented music programs did little to increase the musicality and musical appreciation of the individual musician” (Mark & Gary, 1999, p. 361-362).

Therefore, the principal aim of implementing CM in the public schools was to provide students a broadened music education that went beyond mere technique development. Instructional techniques employed to meet this end revolved around performing, analyzing (describing/listening/evaluating), and composing. These activities utilized the compositions being rehearsed for performance to provide students with a deeper understanding of musical concepts. Hence, the performance ensemble was viewed as a laboratory setting in which musical concepts could be explored actively to enhance learning. This emphasis on conceptual learning was key to the concept of comprehensive musicianship and served to accomplish one of its primary objectives: the transference of knowledge to new settings.

Although performance still was considered vital, it was regarded on an equal footing with other activities (e.g., teacher guided listening, analyzing, evaluating, and composing). Thus, Comprehensive Musicianship, and the various curricula that evolved from it (e.g., The Hawaii Music Curriculum Project, The Manhattanville Music Curriculum Program, The Wisconsin Comprehensive Musicianship Through Performance Project), emphasized an integrated approach to learning in a performance setting that focused on a balanced curriculum. The trend to emphasize this balance between
performing, creating, and responding to music has continued to permeate educational thought since. This is evidenced most recently in the National Standards for Arts Education. Appearing in 1994, the Standards reflect and reinforce Comprehensive Musicianship theory that had been circulating for well over three decades. Indeed, much of the language that was used to describe instructional techniques and strategies in the Comprehensive Musicianship project is paralleled in the language of the Standards:

These Standards assume that students in all grades will be actively involved in comprehensive, sequential programs that include creating, performing, and producing on the one hand, and study, analysis, and reflection on the other. Both kinds of activities are indispensable elements of a well-rounded education in the arts (MENC, 1994, p. 15-16).

In regard to the performing ensemble, the Standards emphasize that:

Every course in music, including performance courses [italics added], should provide instruction in creating, performing, listening to, and analyzing music... (MENC, 1994, p. 59).

Thus, the Standards, in utilizing the same premises that originated decades before, have solidified that music education in performance should strike a balance between the development of performance skills and the development of students' ability to listen, analyze, and comprehend music. It is believed that, in teaching toward these Standards, educators will provide students with the fundamentals necessary to lead meaningful musical lives:
Through singing, playing instruments, and composing, students can express themselves creatively, while a knowledge of notation and performance traditions enables them to learn new music independently throughout their lives. Skills in analysis, evaluation, and synthesis are important because they enable students to recognize and pursue excellence in their musical experiences and to understand and enrich their environment... The ability to listen with understanding is essential...(MENC, 1994, p. 59).

Evident in the above statements is a commitment to develop students' overall musicianship in regard to both performing and listening activities. Ideally then, performing ensemble classes should be providing a balanced music education that focuses on the production and perception of music.

Although sometimes viewed as mutually exclusive, many argue that these generative and receptive behaviors are interdependent constructs which mutually enhance one another. In other words, by learning to perform with proper tone quality, proper intonation, proper expression, rhythmic accuracy, etc., students will be better prepared to perceive these qualities in other performances, and, thus, be able to discriminate and evaluate successfully in regard to new musical experiences. Thus, by building performance skills, the perception of musical nuance (e.g., expressiveness, tempo fluctuations/rubato, dynamics, tonality, etc.) increases simultaneously.

Reimer (1989) claims that students who are involved in performance programs are developing a heightened perception of musical nuance by the intensive study of specific musical works. He insists that performance programs provide opportunities for in depth examination of musical works, and, as a consequence, “musical perception can be commensurate with the
detailed study performance requires" (Reimer, 1989, p. 202). Furthermore, "a
deep probing of perceptual complexities and nuances [is] expected" in the
performance class, thus allowing students to cultivate their musical

Others have suggested that the performing experience itself positively
affects musical perception. For instance, Kivy (1991) suggests that learning to
make music allows one to perceive what others cannot:

...not only is the performing of music, in ensemble, a deep and
rich communal experience; it enriches as well in ways hard to
convey to the nonperformer the experience of musical listening.
It literally makes one able to hear what to others is inaudible (p.
90).

Elliot (1995) contends that making music allows students to be more
cognizant of musical nuance. Although listening to music seems to be an
automatic part of everyday life, "meaningful" listening (that is actively
engaging in what Elliot terms as "listening-for" rather than "listening-to")
requires high levels of attention (Elliot, 1995). The performing ensemble
bolsters and enhances this attention by "plac[ing] the student-as-listener
inside the musical works he or she is endeavoring to learn" (Elliot, 1995, p.
99). By doing so, the musical work and style it represents can be known from
the inside-out, thus allowing generalizability to new musical situations.
Based on Elliot's theories then, the only way to increase perceptual
understanding in students is to focus on making them competent, proficient,
and artistic musicians/performers. For Elliot, perceiving music is contingent
upon making music.
It is important to note that the commonality proposed in these theories centers on the assumed connection between the production of music and the perception of musical nuance. That is, by participating in generative tasks (i.e., performing activities), students will become more competent in receptive activities (e.g., identifying musical nuance). If this assumption is true, then those who have participated in performance courses should possess a heightened perception of musical nuance compared to those who have not. Also, the degree of enhanced musical perception should be contingent on the amount and degree of experience in performance. That is, the more expertise (e.g., years) of experience one has in performance, the more perceptive one’s abilities. Unfortunately, research indicates that this assumption of transfer is far from conclusive.

Need

Research investigating the link between performance and perception is somewhat mixed. The National Assessment of Educational Progress (NAEP), for example, targeted music as a content area in 1997 in order to assess what students “know and can do” in music (National Assessment Governing Board, 1994, p. i). This study assessed 8th grade students’ musical ability in performance and perceptive tasks, and compared students enrolled in music classes three or four times a week, once or twice a week, less than once a week, and those not enrolled in music at all. Based on the findings of the study, NAEP concluded that “there appear[s] to be no consistent relationship between achievement [in music]... [and] frequency of instruction, [or] curriculum availability” (NCES, 1999). Although not reported as statistically
significant, some of the report's findings deserve further attention. With respect to those enrolled in performance classes, those enrolled in choir everyday achieved a higher perceiving score than those not enrolled in music at all. Likewise, those enrolled in instrumental ensembles scored higher than those not enrolled. In light of this evidence, it is difficult to conclude if performing is having a consequential effect on perceptive tasks.

In another study, DeTurk (1988) investigated performing ensembles influence on students' critical thinking about music. Specifically, he examined the relationship between experience in performing music and the development of critical thinking skills about music. DeTurk concluded that there appears only to be a weak relationship between years of experience performing music and level of critical thinking about music. In post hoc analyzes, he found that at least six years of experience in performing was required before a significant difference in critical thinking was evident.

Other studies have found weak or non-significant links between performing and perceiving. As mentioned above, Mercer (1972), Duerksen (1968), and Benner (1972) indicated that participation in performance activities had only minor effects on musical perception. Furthermore, Killian (1991) found that there was no significant relationship between sight-singing ability (production) and error detection skills (perception). Similarly, Apfelstadt (1984) and Geringer (1983) found non-significant relationships between melodic perception and performing (i.e., singing) ability.

While these studies have suggested that there is a weak or non-existent connection between performing and perceptive abilities, other studies have concluded the opposite, suggesting that performing experience directly affects the perception of musical nuance. Morrongiello, Roes, and Donnelly (1989),
for example, in studying children's perception of musical patterns, concluded that children with musical training in instrumental performance showed enhanced sensitivity (perception) to specific melodic features (i.e., individual frequencies). Halpern (1984) and Cuddy and Cohen (1976) reported similar findings in adult populations with regard to minor/major key change recognition and transposition discrimination. Likewise, Sheldon and Gregory (1997) found that the perception of tempo modulation was affected by level of performing expertise (i.e., number of years of musical experience). These findings support other research that has focused on the qualitative differences between musicians' and non-musicians' perception of musical nuance (Bartlett & Dowling, 1980; Geringer & Madsen, 1981; Madsen & Geringer, 1990; Ramsey, 1983; Wapnick & Rosenquist, 1991). (For a thorough exegesis of the pertinent literature see Chapter 2).

In light of the extant research, the influence of performing on the perception of musical nuance is still in doubt. With respect to music education curricula, the influence of the performing ensemble on subsequent perceptual and productive abilities also is questionable. Obviously, further investigation in this area is warranted.
Statement of the Problem

The present study was designed to investigate if the cultivation of generative and receptive behaviors is taking place through the performing ensemble curriculum. Specifically, the degree to which the performing ensemble experience effects the production and perception of musical nuance was considered.

For the purposes of this study, the production and perception of musical expression was the principal musical nuance investigated. As expression involves many musical subtleties working in concert (e.g., phrasing, articulation, tempo, dynamics, timbre), it is a strong starting point in investigating the performance and perception of musical nuance. That is, it is assumed that these components working together are more salient than each working in isolation. Furthermore, as Kohut (1973) has indicated in his text on performance pedagogy, "the essence of musicianship is one's sensitivity to the expressive qualities of music, to the nuances of phrasing, interpretation, and style- in brief, all of those elements which make musical performance an art as well as a skill" (p. 6). Thus, in studying expressiveness, the present study attempted to evaluate how successful performance ensembles are in educating to this end. Additionally, it has been argued by others that by investigating expression, overall musical understanding is assessed as well (Jackendoff, 1991; Lerdahl & Jackendoff, 1983; Drake & Palmer, 1993, Todd, 1989). If so, then studying competence in expressive behavior (both generative and receptive) could validate the success of the performance ensemble in teaching toward comprehensive understanding. With respect to selected Standards (e.g., six, listening to, analyzing, and
describing; and, seven, evaluating music and musical performance) which pertain specifically to students’ ability to perceive musical nuance, the effect of performing ensembles in fulfilling these standards would be addressed as well.

Purpose

The purpose of this study was to examine the effects of school performing ensemble experience (i.e., band, choir, or orchestra participation) on students’ ability to perceive expression in music and to perform a musical passage expressively using a minimum skill musical device. Because musical expression is contingent to a large degree on the accuracy of the musical performance, the ability to perform a musical passage accurately was examined as well.

Two sample groups of participants with differing amounts of school performing ensemble experience were compared. The first group consisted of participants with no musical training, while the second group consisted of those who had a minimum of one year of high school performing ensemble experience. Specific research questions that were addressed are as follows:

1. Is there a significant difference between subjects who have had school performing ensemble experience and subjects who have not in the ability to perform a musical passage with expression?

2. Is there a significant difference between subjects who have had school performing ensemble experience and subjects who have not in the ability to perform a musical passage accurately?
3. Does the number of years participating in a school performing ensemble affect the ability to perform a musical passage with expression?

4. Does the number of years participating in a school performing ensemble affect the ability to perform a musical passage accurately?

5. Is there a significant difference between subjects who have had school performing ensemble experience and subjects who have not in their ratings of perceived expression in a musical passage?

6. Is there a significant difference between subjects who have had school performing ensemble experience and subjects who have not in their ratings of perceived accuracy in a musical passage?

7. Does the number of years participating in a school performing ensemble affect ratings of perceived expression in musical passage?

8. Does the number of years participating in a school performing ensemble affect ratings of perceived accuracy in a musical passage?

Definition of Terms

The following terms, used in the Introduction and throughout this study, will be defined as follows:

Musician: Participants in the present study were labeled as musicians if they had at least one year of high school performing ensemble participation. The musician group was limited to those persons whose formal musical training was through school performing ensembles exclusively. That is, these subjects had no formal musical training beyond the school performing ensemble. Furthermore, because of the findings of Flowers, Dregalla, and
Jackson (1993), keyboardists were eliminated from the musician group (see Chapter 2 for details). Number of years of participation in a school ensemble was calculated by adding together the number of years enrolled in elementary, middle school and high school performance ensemble courses. Thus, a subject with two years of high school and two years of middle school was labeled as having four years of participation.

Nonmusician: Participants in the present study were labeled as nonmusicians if they had no previous formal music instruction. In defining the population thus, it is acknowledged that most persons falling in this category would have at least some musical background. That is, most would have had experience in elementary school music programs and extensive contact with music through cultural contexts. The designation of nonmusician within the context of this study indicates that the person has never played a musical instrument or formally performed in an ensemble beyond these basic experiences.

Performance classes/courses: Courses in which teaching performance skills/concepts and providing students with opportunities to perform represent the primary curriculum of the class. Examples include choir, band, and orchestra.
Expression: The Harvard Dictionary of Music (1964) defines expression as the:

part of music which cannot be indicated by notes or, in its highest manifestation, by any symbol or sign whatsoever. It includes all the nuances of tempo, dynamics, phasing, accent, touch, bowing, etc. by which the mere combination and succession of pitch-time-values is transformed into a living organism. [Expression] may be said to represent the creative contribution of the performer...The ideal performer is the one who succeeds in bestowing upon the composition a personal and original expression within the stylistic frame of the work... (p. 250-251).

Expressive performance: Any performance in which the performer actively manipulates dynamic, rhythmic, and articulative nuances to correspond with structural elements within the composition to render an individual interpretation.

Unexpressive/mechanical performance: Any performance that is devoid of dynamic, rhythmic, and articulative (i.e., expressive) fluctuations. Performances of this nature also would not mark phrase endings or structural events.

Dynamic variation: Increases or decreases in amplitude (loudness) in a performance deliberately manipulated by the performer in an effort to render an expressive performance.
Tempo variation/Rubato: Increases or decreases in the temporal aspects (speed) of a performance deliberately manipulated by the performer in an effort to render an expressive performance.

Musical Accuracy: For the purposes of this study, musical accuracy was defined as the ability to perform with precise rhythms. As the *Instant Pleasure* software used in this study (described later) generated the correct notes for each performance, note accuracy was not a concern for subjects.

**Limitations**

It is beyond the scope of this study to address emotion or intended emotion as a concomitant of musical expression. Many definitions of musical expression tend to equate emotion as a product of or synonymous to expressive qualities. *The New Harvard Dictionary of Music* (1986), for instance, holds as one of its definitions of expression: “the representation or conveying of something, usually something beyond the work itself such as a nonmusical idea or emotion” (p. 295). This definition, oftentimes regarded as referentialism, is not addressed in the present study.

**Overview of Remaining Chapters**

Chapter 2: Review of Literature

Chapter 2 provides an overview of empirical studies in the perception and production of musical nuance. Specific regard is given to those studies that examine the production and perception of musical expression, and the
influence of training and musical experience on these behaviors. The chapter is intended to clarify further the scope, purpose and significance of the present investigation and its relationship to the extant body of research in this area.

Chapter 3: Methods and Procedures

Chapter 3 describes in detail the procedures used in this study. The chapter begins by restating the questions under investigation and then describes the population selected for study, the selection of research participants from this population, and the specific tasks that participants were asked to perform. Also, included is a detailed description of the experimental apparatus employed to collect data. The chapter concludes with a description of data analysis procedures.

Chapter 4: Results

Chapter 4 presents the findings of the current investigation. In this chapter, data pertinent to each research question is presented separately.

Chapter 5: Summary, Discussion, Conclusions, and Recommendations for Further Research

Chapter 5 begins with a summation of the data presented in Chapter 4. The relevance of the specific findings is discussed and related to previous research in this area. Concluding remarks state the implications of the present study for music teaching and learning, and suggests additional research directions that will continue to inform the understanding of the phenomena.
CHAPTER 2

REVIEW OF LITERATURE

The extant literature regarding the performance and perception of music is rich in history. For almost a century, empirical research concerning the performance of music can be evidenced. Sears (1902), an early pioneer in the field, devised a mechanism for measuring note lengths in live organ performance. His apparatus for measurement consisted of steel pins connected to the back end of organ keys. When a particular key was depressed, the attached pin dipped into a mercury cup and made an electrical contact that allowed for the recording of note durations, accurate to 0.01 seconds. Following Sears's lead, Morton (1920) and Hartmann (1932) investigated musical performance; the former using an apparatus similar to that of Sears, the latter using player piano rolls. Each of these studies was concerned with rhythmic variation found in live performances. By measuring the duration of notes, these researchers could discern deviations from exact notational ratios (e.g., 3:1 = dotted half note to quarter note; 2:1 = quarter note to eighth note, etc.), as well as note overlappings, and note asynchronies. It was believed that these deviations were not random, but, in fact, were intentional deviations used to communicate structure in music and individual interpretation.
Likewise, empirical research concerning the perception of music also is rich in history. Schoen (1928), for instance, concluded that the affective state suggested by a composition is fairly constant and universal, and that interpretive activities by listeners in assigning meaning to music is a common occurrence in the listening process. Building on Schoen’s research, Hevner (1936) studied the affective value of expressiveness in music by creating an adjective matrix which listeners would use to classify their responses to musical stimuli. Of primary importance in her investigation was the use of musical structure in communicating expression (i.e., major/minor mode, melodic contour, rhythm, harmonic structure). She also was one of the first to note that musical training is integral to the perception of musical meaning:

"[Listeners’] attitude and their success in apprehending the music vary with their training and sophistication in music matters...and [with] their backgrounds in musical and non-musical experience (Hevner, 1935, p. 187)."

The early period of research in the performance and perception of music culminated with the work of Carl Seashore. Seashore and his colleagues introduced new technologies in the study of musical performance that allowed for the measurement of both duration and intensity. These studies (Greene, 1937; Henderson, 1937; C. Seashore 1932, 1936, 1938; H. G. Seashore, 1932, 1937; Skinner & Seashore, 1937; Small, 1937) produced voluminous data that established a solid foundation for the study of performance. Seashore’s studies were able to confirm that inter-individual performances did, in fact, contain many common features (e.g., retardations at the end of phrases, accelerating at the beginning of phrases, asynchronous
nature of chords, etc.). In addition, these studies indicated that intra-
individual consistency from performance to performance was usually high.

Most importantly, Seashore was one of the first to acknowledge the
interrelationship of performance and perception as mutually supportive and
informative precepts. In discussing the findings of his studies in
performance, he commented:

The psychophysical relations between the performer and the
listener must be worked out; the data presented will contribute
to such studies and also will depend for their final interpretation
upon such investigations (Seashore, 1937, p. 118).

By qualifying his performance studies in this manner, Seashore recognized
that the study of perception would provide insight, clarification, and
confirmation of performance activities. Thus, the study of performance alone
is incomplete without the study of perception. In regard to perception, it is
important to note that Seashore and his colleagues shared Hevner's view that
perception was contingent upon training and experience. Specifically,
Seashore (1938) commented that "training...can greatly increase the functional
scope of these capacities" (p. 3).

Since Seashore and Hevner, research in the production and perception
of musical nuance has grown extensively. With the advent of new
technologies, most importantly the Musical Instrument Digital Interface
(MIDI), research in this field has blossomed, particularly in the last 15 years.
Following the lead of Seashore and Hevner, studies have continued to
investigate the performance and perception of musical nuance as well as the
effect of training on these abilities. Unfortunately, as most of these studies
concentrate on expert musicians or non-musicians, there is a paucity of
research centering on the extent and depth of musical training, especially concerning the influence of school music curricula on these abilities. Furthermore, although many studies do acknowledge and investigate the link between performing and perceiving music, few studies have investigated the impact that performance ensemble participation has on the perception of musical expression.

The remainder of this chapter presents an overview of the studies most closely related to the present investigation. The chapter is divided into three sections. The first section examines studies that explicate the nature of musical expression. Specific attention is given to the communication of expression through performance practices. Section two expands on the first section by presenting studies that investigate musicians' and non-musicians' perception of the salient features inherent to musical expression communicated by performance. Section three presents research that is most similar to the present investigation. It includes a survey of the research that explores the impact of curriculum on musical perception, and a discussion of two studies using the Instant Pleasure software program. It is important to note that these sections are interrelated and will overlap to some extent. It is only by viewing the research as a whole that a thorough understanding of the phenomena can be gleaned.
Studies in the Performance of Musical Expression

Researchers have operationally defined expression as the deliberate deviation from exact notation. Seashore (1938) in his studies at the University of Iowa found that musicians systematically and purposefully deviate from notation in terms of tempo, duration, intonation, intensity, articulation, and vibrato. Although he found that musicians do not play two identical notes in the score in the same manner, the deviations implemented by performers are, in fact, systematic and oftentimes predeterminate. Researchers since Seashore have followed this basic operational definition. Sloboda (1994), in discussing the extant research in musical expression, provides the following as a definition of expression:

Timing, loudness, tone quality, and intonation can be substantially varied, even within the same measure. Such variations together constitute what we usually call "expression" (Sloboda, 1994, p. 154).

He goes on to indicate that these expressive variations are not random or arbitrary, but are systematic in nature. That is, performers cannot vary the nuance of any particular note or notes at random; "it matters which notes we vary and how" (Sloboda, 1994, p. 154). The logical consequent of this is that there are certain parameters in music that constrain musical expression. An obvious factor that would constrain expression would be the stylistic and cultural features of the piece performed. That is, tradition, acculturation, and musical style all have a direct impact on the amount and degree of expression employed. In addition to these concomitants, many have theorized that

To appreciate this fully, consider Lerdahl and Jackendoff’s generative theory of tonal music. This theory suggests that music involves the encoding and decoding of structural elements found within the music to communicate a cogent musical hierarchy. Pitch and rhythm, therefore, constitute the transformational devices with which music is communicated to form “an explicit formal musical grammar that models the listener’s connection between the presented musical surface of a piece and the structures he attributes to the piece” (Lerdahl & Jackendoff, 1983, p. 3). The role of expression within this system is a generative one that functions to convey aspects of the musical structure such as boundaries and segments that allow the listener to form a hierarchal conception of the piece. This is important because music is a temporal art that must be processed over time. Segmentation of musical events occurs, then, in order to allow for cogent processing of the received temporal stimuli. In this regard, musical expression allows for more salient musical processing because it assists the listener in assigning/generating these structural properties.

Much research into the performance of musical expression has centered on examining the theories of expression discussed above. As such, many studies have been interested in the systematic variations of timing, intensity, and articulation that impact musical expression, as well as how these systematic variations correspond to and communicate/generate musical structure. In investigating the former, researchers have been particularly interested in separating random or unintentional fluctuations in timing,
intensity, and articulation from those that are deliberate and intentional. One way to examine this is to measure the expressive parameters employed by a particular performer in several performances of the same work. By utilizing this method, unintentional errors are randomly distributed throughout the various performances, thus allowing the variations that appear consistently in each performance to be regarded as intentional. Using this method, Shaffer (1984), examined beat durations in three performances of a Chopin etude performed by the same pianist. The performances, although separated by as much as a year, reflected considerable consistency. Not only were global durational features consistent (i.e., major slowing at particular measures), beat durations within measures were consistent as well. Other studies, using this same method, also have demonstrated that professional musicians perform with exceptional consistency (Repp, 1996; Seashore, 1938; Shaffer, 1981; Shaffer, Clarke, & Todd, 1985). Although these studies show consistent expressive intent, they do not examine the degree to which deviations from the score are contingent on structural devices.

To investigate the supposed effect of structure on musical intent, Sloboda (1983) carried out an experiment in which pianists with differing levels of experience (12 - 40 years) performed several different melodies at sight. Each melody was paired with a “twin” that shifted the bar lines one note along the original melody. The twin melodies, subsequently, were randomized and separated by other melodies so that the pianists would not recognize them as the same. Although the melodies were identical in every way except metrical grouping, pianists performed each melody quite differently. These differences corresponded to metric structure— that is, the first note of a measure was played significantly louder, longer, and more
legato than the notes on either side; however, these parameters differed by amount and degree depending on the pianist. It was discovered that pianists with more experience employed a greater number of metrically related parameter variations than did lesser experienced pianists. Specifically, the pianists with more experience employed greater contrast in expressive variation and were more consistent in their execution (i.e., notes beginning the measure or half measure were consistently played louder, longer, and more legato). In contrast, lesser experienced pianists did not employ as great a range of expressive variation, and oftentimes deviated from playing notes beginning the measure or half measure louder, longer, and more legato. This finding supported Sloboda’s theory that “expert performance is governed by underlying structure to a greater extent than that of novices” (Sloboda, 1985, p. 274).

Additionally, in the same study (Sloboda, 1983), a listening task was employed to determine which performances communicated metrical structure more readily. Each pianist’s performance was played to a group of musicians who, in turn, selected the appropriate barring of the melody. Sloboda found that listeners were better able to identify correctly the metric structure in the more experienced pianists’ performances. From this, Sloboda concluded that increasing the number of parameter variations to correspond to metrical structure positively effected meter identification.

In a follow up study aimed at exploring these parameter variations more deeply, Sloboda (1985) attempted to elicit the salient expressive parameters employed in communicating metric structure. Having identified these expressive variants in his previous study, Sloboda employed two pianists, one experienced and one novice, and also used a set of computer
generated melodies that varied the amount and number of expressive parameters employed in performance (i.e., duration, articulation, and loudness) to examine metric communication. Conclusions paralleled that of the first study. The expert pianist was able to employ expressive parameters that communicated metrical structure more consistently than the lesser experienced pianist. Specifically, the expert performer used a greater range of loudness and was more skilled in his use of articulation. The experienced pianist also tended to link loudness and articulation cues together, whereas the lesser experienced pianist did not. In regard to the computer generated performances, Sloboda found that variations in duration were not able to communicate metrical structure on their own, whereas changes in loudness alone were highly effective in communicating metrical structure.

From these studies, Sloboda (1985) concluded that “skilled music performance is intimately mediated by mental representations of metrical structure,” and that “these representations mark boundaries of principal metrical subdivisions...for expressive treatment” (p. 291). Skilled musicians, then, manipulate expressive devices to communicate more clearly a representation of musical structure. Furthermore, the communicated representations are not necessarily premeditated, but can be generated or triggered during performance.

Others have expanded on this by examining how expression is contingent upon the performer’s knowledge of musical structure. Clarke and Baker-Short (1987), for example, had musicians imitate the rubato of short musical phrases in which the relationship between structure and expression was varied. The relationship between rubato and tonal structure was manipulated so that four variations occurred. The expression (rubato in this
case) in some variations would be consistent with the phrase’s tonal structure, whereas in other variations the expression would conflict directly with tonal structure. The subjects, four pianists with a minimum of 10 years experience, listened to each variation three consecutive times and then were asked to reproduce the melody on the piano. Results indicated that the pianists were better able to reproduce rubato when it was linked to the phrase’s underlying tonal structure. Phrases in conflict with underlying structural properties were not performed as successfully. Clarke and Baker-Short concluded from this that “convincing expressive performances can only be reliably achieved when the underlying musical structure has been understood and assimilated” (p. 58).

This conclusion by Clarke and Baker-Short reinforced Clarke’s previous work that examined the interaction between musical structure and performance tempo/rhythm. In this study, Clarke (1982) analyzed performances of Satie’s Vexations for consistent deviations from metrical timing. He found that these deviations produced timing curves in which specific boundaries could be evidenced. Comparing these boundaries with the structure of the piece, Clarke found that these boundaries corresponded significantly to positions in the musical structure where formal breaks could be identified.

In light of these studies, Todd (1985, 1989) proposed a hierarchical model of expression as it relates to the performance of rubato. The model suggested that tonal music is most similar to speech in terms of hierarchical structuring and grouping principles. Taking cues from prosodic and phonological theories and research, Todd suggested that performers utilize grouping structure, metrical structure, time-span reduction, and prolongation
reduction in communicating musical ideas. These components assist the listener in processing an entire piece of music into cogent “pieces.” His model focused specifically on phrase final-lengthening in performance, and suggested that performers utilize this form of rubato “to reflect some underlying structure abstracted from the musical surface” (p. 34). To test this model, Todd (1985) examined the timing parameters employed in performances of piano compositions in the Western art music tradition. He found that ends of phrases played a significant role in determining the pianist’s use of expressive timing. Ends of phrases were lengthened by performers to reflect hierarchical structuring at points of relative stability. That is, by using rubato, “the performer selects a structure from the set of possible structures and elucidates this to the listener” (p. 49). These findings supported the tested theory that knowledge of musical structure is essential to the effective use of expression.

While Todd (1985, 1989), Clarke (1982), and Clarke and Baker-Short (1987) have examined the effects of temporal changes in communicating musical structure, others have examined temporal asynchronies. Palmer (1996b), for instance, investigated how performers utilize temporal asynchronies to communicate melodic structure in piano performance. In a series of experiments designed to examine the nature of melodic communication, she found that pianists used consistent note onset asynchronies when performing melodic lines with accompaniment. Specific findings indicated that the melody line was played slightly earlier than notated simultaneities. That is, the melody preceded other events in performances of both familiar and unfamiliar music. Furthermore, the study indicated that pianists with more experience utilized longer melody leads in
both familiar and unfamiliar music. It was found, however, that lesser skilled pianists began to utilize melody leads as their practice time with unfamiliar music increased. Palmer suggested that this may indicate that the "pianists' level of experience influence[s] expressive performance of unfamiliar music," and that more experienced musicians may include expressive devices in performance sooner than musicians with lesser experience (p. 39). Also worthy of mention is that listeners engaged in perceptive tasks were better able to identify the melody when melody leads were present, implying that the communication of musical structure is assisted by the variation of musical nuances.

Palmer's use of familiar and unfamiliar music deserves note. This aspect was included in her study to determine how much flexibility performers have with expressive features, and if this flexibility is contingent upon expertise. Earlier studies (Drake & Palmer, 1993; Palmer, 1989; Seashore, 1938) have found that even when expert musicians attempt to perform familiar (i.e., practiced or rehearsed) pieces mechanically, expressive gestures are oftentimes residuary. This suggests that well learned behaviors such as expression are not as malleable as others have suggested (Sloboda, 1994; Todd, 1985), but instead are ingrained into performance practice. To assess this, Palmer asked pianists to perform familiar and unfamiliar excerpts both musically and mechanically. Her results indicated that there was a significant difference in melody leads between musical and mechanical performances in both familiar and unfamiliar excerpts, but this difference was contingent on level of performing experience. That is, more experienced pianists employed more melody lead contrast between musical and mechanical performances than did lesser experienced pianists regardless of the familiarity of the excerpt.
Several conclusions relevant to the present study can be made from Palmer’s experiments. First, level of experience in music does seem to play a significant role in the performance of musical nuance. As evidenced, not only did level of experience effect the performance of familiar music, but also the performance of unfamiliar music. Second, expressive performance gestures appear to be more ingrained in expert musicians and therefore need to be controlled for by employing the use of unfamiliar material. As Palmer (1996b) suggested, “extensive practice of particular performances may make expressive gestures more difficult to suppress...[therefore] expressive gestures may be more subject to change in situations in which their specific implementation is not yet well-learned” (p. 36, 54). Third, pianists seem to move easily between various domains of expression:

Pianists were able to perform without expression even when the performance was well-practiced (Experiment 1), perform expressively when the music was unknown (Experiment 2), and alter the particular melodic interpretation being expressed (Experiment 3) (Palmer, 1996b, p. 53).

Finally, experts may own particular skills (e.g., knowledge of style/period) that allow for quicker interpretation of musical stimuli.

Other researchers also have examined expressive devices in musical performance that mark musical structure. Drake and Palmer (1993) analyzed systematic performance variations (i.e., intensity, inter-onset timing, and articulation) found in three accent structures: rhythmic grouping, melodic, and metric accent structures. Results suggested that musicians use systematic variations in timing, intensity, and articulation to emphasize these accent structures. Dynamics, in particular, were found to mark melodic contour and
emphasize stronger metric locations. Similarly, Todd (1992) found that
dynamics are often used to indicate phrase structure, articulate group
boundaries, and signify important structural events. The same premises are
found in regard to articulation. Palmer (1989) found that pianists employ
more legato style when asked to play expressively versus mechanically.
Likewise, Clarke (1984) found that there appears to be a weak, but positive
relationship between structural significance and legato articulation.

Clearly, many expressive devices are possible for manipulation in
musical performance. As Sloboda (1988) has indicated, however, these
variations cannot be randomly applied, but are dependent on musical
structure. Clarke (1988) sums:

Within each of the three expressive parameters (timing,
dynamic, articulation), expressive gestures can perform a
number of different functions: these include altering the relative
proportions of events within a rhythmic group, indicating the
position of a group boundary, marking a metrical accent, and
creating an expressive gradient towards a focal point (that is, a
pattern of directed motion). As a result, expressive gestures are
functionally ambiguous, in the sense that they specify a number
of alternative interpretations. It is only through interactions
with the underlying musical structure that these ambiguities are
resolved (p. 13).

The studies discussed above all reflect a particular method of
investigation into the phenomenon of musical expression. As these studies
rely on measuring expressive devices employed in performances, they
typically are referred to as analysis-by-measurement studies. The general
principle, is to determine the consistent musical nuances elicited in repeat
performances and then to generalize these findings to specific rules concerning the nature of expression. Many of these rules center on structural features of the performed music.

Another approach to investigating musical expression is the analysis-by-synthesis approach, employed, most notably, by Sundberg, Fryden, and Askenfelt (1983). Analysis-by-synthesis studies compose or synthesize musical performances based on the advice of expert musicians. For instance, an expert musician might suggest certain changes within the synthesized performance to enhance musical expression. Dynamics, articulations, and durations thus can be altered in order to form an ideal expressive rendition of a piece. This ideal version then can be analyzed for specific rules that occur in regard to expressive parameters. From this, experimental studies testing these rules can be generated to examine the overall salience of each expressive device and/or devices in combination.

Sundberg, Fryden, and Askenfelt (1983) used this procedure to generate rules for expressive musical performance. Some of the generated rules are applicable to relatively short time spans (e.g., three to four notes) such as: the higher the frequency, the louder the amplitude; short note values are played shorter to heighten durational contrast; shortening the starting tone in rising intervals; etc. Other generated rules apply to longer time spans (e.g., phrases, movements, etc.) such as: dynamics and rubato are contingent upon structural harmonic charge; and phrases are marked with final phrase lengthening followed by a short pause. It is important to note that all generated rules were found to correlate significantly with aspects of the underlying musical structure.
In order to test the accuracy and salience of these generated rules, Thompson, Sundberg, Friberg, and Fryden (1989) allowed listeners to evaluate the effectiveness of each rule(s) in context. Several melodies were extracted from Western art music and constructed into five variations: 1) mechanically; 2) applying only the rule being tested; 3) applying the opposite of the rule being tested; 4) applying three rules (all but the rule being tested); and 5) applying all four rules. In testing the rules relating to short time spans, Thompson et al. found that applying the rules in isolation increased significantly the expressivity ratings of the performances. Furthermore, it was found that the application of all the rules at once produced significantly higher ratings than when applying three of the rules simultaneously. In testing the rules relating to more global aspects of the music (longer time spans), results were not as clear. Most significant was the interaction between the musical excerpt and the rule applied. This suggests that these type of rules “may depend on knowledge of high levels of musical structure” which were not captured in a global application of these rules” (Thompson et al., 1989, p. 79). Thompson et al. concluded that more investigation was warranted in order to elucidate how subtle changes in musical structure effects the applicability of “longer time span” rules.

The analysis-by-synthesis approach by Sundberg et al. and Thompson et al. provides converging evidence for the theory that musical expression is contingent upon musical structure. In these studies, it was clear that higher ratings were given to performances when a greater number of rules applied. As these rules relate specifically to structural points within the music, this suggests that “better performances will require a greater amount of musical
knowledge” (Thompson et al., 1989, p. 79). This theory is consistent with the studies previously reviewed: knowledge of underlying musical structure is requisite to performing with exemplary musical expression.

In general, these studies indicate the deliberate nature of musical expression. Performers are shown to systematically deviate from strict notational practices in order to achieve more musical performances. These deviations have been shown to coincide with important structural events found within the musical context. A performer’s knowledge of these structural events will assist in determining where and to what degree variations in musical nuance will be implemented. It was suggested that a more thorough understanding of the underlying musical structure will positively affect the performer’s use of expressive devices. In the next section, the use of these expressive devices to positively affect the perception of musical nuance will be discussed. This will be followed by a review of studies that examine more closely the effect that musical training has on musical perception and production.

Studies in the Perception of Musical Expression

Researchers have focused on the perception of musical expression in much the same way that they have focused on the performance of musical expression. As such, research in the area of the perception of musical expression has been concerned with the communication of expression to the listener through the performer’s manipulation of various musical nuances (e.g., loudness, timing, articulation, rubato, etc.) As in many of the performance studies discussed above, researchers have been interested in
how the systematic variations of these musical nuances guide perceptive processes. Put another way, these studies aim to discover if the variations of musical nuance found in musical performance communicate effectively musical structure and expression to the listener.

Many researchers have studied the communication of musical expression by determining if changes in musical nuance such as duration, dynamics, and articulation are perceived by listeners. In investigating listeners' perception of expressive timing, Clarke (1989) examined the ability of listeners to identify durational changes in music. Subjects were instructed to listen to three different versions of a composed melody and determine if the melodies were mechanical or contained expressive devices (i.e., deviations from strict metronomic performance). He found that subjects were able to perceive expressive timing changes in performances as well as note lengthenings as minute as 20 milliseconds in a context of notes lasting 200 to 400 milliseconds. Clarke concluded that timing variations are a salient means by which to communicate structural properties in music. Similarly, Nakamura (1987) studied listeners' ability to perceive performers' intended dynamic changes in music. The researcher had professional musicians record two movements of a Baroque sonata inserting the dynamics on their own. At the conclusion of the performance, the musicians were asked to indicate in the score both the location of the dynamics and the degree of intensity used. The performances were played to listeners who indicated in a blank score where they perceived dynamic changes and what intensity level corresponded to each dynamic. Nakamura found that, in general, performers were able to convey their dynamic intent to listeners accurately. Listeners perceived crescendos more easily than decrescendos, however, and small
changes in dynamics oftentimes were not recognized. Specifically, crescendos that increased in intensity less than 2 dB were identified only as likely as that which would be identified by chance.

In a more gestalt approach, Crist (1996) examined listeners' abilities to distinguish between expressive and mechanical performances. In this study, four musicians were asked to play a musical excerpt twice, once expressively and once mechanically. The performances were recorded and played for subjects (musicians) in two listening tasks. In the first task subjects were asked to identify the performance as either expressive or mechanical. In the second task, each expressive performance was paired with an unexpressive performance and subjects were asked to indicate which of the two performances was expressive. Crist found that listeners were able to accurately discriminate between expressive and mechanical performances.

In a follow up study, Crist (1998) investigated listeners' perception of tempo and dynamic variations and its effect on listeners' ability to identify an expressive performance. Most subjects were instrumental music education majors. Crist found that tempo and dynamic changes were readily identified by subjects, thus lending support to the findings of Clarke (1989) and Nakamura (1987). Furthermore, Crist found that subjects identified more easily changes in tempo than in dynamics. He suggested that this may indicate that performers of the musical excerpts were more proficient in communicating tempo variations over dynamic variations or that the subjects were better able to perceive tempo variations. In regard to the effect of tempo and dynamics in the communication of musical expression, Crist found that subjects rated the excerpts with both tempo and dynamic fluctuations as most expressive. The next highest rankings were given to
performances that contained tempo variations only, followed by dynamic variations only. Performances without dynamic and tempo variations were rated significantly lower than the other performances.

As can be seen, for expressive nuance to be perceived as effective, not only does the intended expressive variation(s) need to be present, it needs to be of sufficient magnitude in order to be detected by the listener. In addition to this, Sloboda (1994) suggests that the expressive variation needs to be interpretable “to ensure that the variation mean[s] something to the listener...at either an intellectual, aesthetic, or affective level” (p. 158). This implies that the intellectual judgment of the listener must be engaged to perceive the occurring expressive variant. As such, the expertise and musical background of the listener comes immediately to the fore in shaping the musical experience- a connotation reflected in the writings of Reimer (1970, 1989, 1995), Kivy (1991), and Elliot (1995) (see Chapter 1). Because these experiences are thought to play such a prominent role in the shaping of listeners’ perceptions, much of the research into the perception of musical expression is dedicated to exploring the effect that expertise has on perceptive abilities.

Before addressing perception as it relates to musical nuance and expression specifically, several studies that explore the effect of training/expertise on listening tasks deserve note. Madsen and Fredrickson (1993), in a replication of Neilsen (1983), examined musicians’ and nonmusicians’ perception of tension in music. Using a continuous response measure (CRDI-the Continuous Response Digital Interface), subjects listened to a recording of Haydn’s Symphony No. 104 and responded continuously to perceived levels of tension in the performance. Results indicated a strong similarity to
Nielsen's findings. In both studies, musicians and nonmusicians reported similar perceptions of tension throughout the performance. In a subsequent study, Madsen, Byrnes, Capperella-Sheldon, and Brittin (1993) investigated musicians' and nonmusicians' aesthetic response to music. The study again employed a continuous response measure manipulated by the subjects to indicate their perceived aesthetic response to music. Subjects listened to five recordings from the Western art music tradition, including the same recording of Haydn's Symphony No. 104 employed in Madsen and Fredrickson (1993). Once again, results indicated strong similarities between musicians and nonmusicians in their response to music. That is, there was no significant difference between musicians and nonmusicians in their perceived aesthetic response. Other studies have continued to confirm this trend in perception of global musical parameters by musicians and/or nonmusicians (Capperella-Sheldon, 1992; Misenhelter & Lynch, 1997).

The findings of these studies imply that the perception of global parameters in music such as tension and aesthetic response may be "universal." Consequently, these types of perceptive tasks may be more a product of acculturation than training in music. Madsen, et al. (1993) elaborate:

Another aspect relates to formal music study. Evidently, nonmusicians...had very much the "same" aesthetic experience as the musicians. While it could be argued that these nonmusicians had a great deal of similar music experiences when compared to their music major counterparts, the additional formal study of music does not seem to have an appreciable gain in aesthetic sensitivity. Perhaps, just growing up in a particular cultural environment provides the necessary ingredients for adequate aesthetic responsiveness...(p. 188).
Although the studies above speak to the similarities between musicians and nonmusicians on listening tasks, other studies have shown that musicians and nonmusicians listen to music in qualitatively different ways. In a study involving musically trained and untrained children between 4 and 6 years of age, Morrongiello, Roes, and Donnelly (1989) found that musically trained children discriminated more accurately among the transformations of an unfamiliar six tone melody. The reason cited for the musicians higher rate of success was a distinct difference between the two groups' focus of attention. Whereas musical children were found to have a higher sensitivity to more specific melodic features (i.e., individual frequencies), children without musical training attended primarily to general pattern features.

Similarly, Madsen and Geringer (1990), in a study involving college students, found that music majors attended to different musical parameters than non-music majors. By using a continuous response measure, subjects were asked to indicate their ongoing focus of attention to either rhythm, dynamics, timbre, melody, or everything. Results indicated that musicians spend the majority of their time listening to melody, then rhythm, dynamics and finally timbre. The nonmusicians listened to dynamics primarily followed by melody, timbre, and everything. A low percentage of time was spent listening to rhythm by nonmusicians. In a follow up study using a different instrument for the dependent measure, Geringer and Madsen (1996) asked musician and nonmusician subjects to listen to a musical selection in its entirety, and then indicate, by assigning percentages, the amount of time spent listening to either melody, dynamics, timbre, rhythm, or everything. The musical selections consisted of two musical exemplars selected to be
representative of each listening category (i.e., melody, dynamics, timbre, rhythm, and everything). Results of the study supported the findings of the previous study—musicians and nonmusicians differed in their focus of attention when listening to music. Specifically, musicians were more consistent in attending to the more salient feature exemplified by each excerpt. Beyond this, timbre was regarded as most salient by musicians, whereas dynamics and melody were regarded as more salient for nonmusicians. In comparing the results of this study with their previous findings, Geringer and Madsen found that musicians were much more consistent than nonmusicians in executing the two tasks. That is, there appeared to be consistent listening patterns in the musician group which would indicate that musical training is affecting perceptual abilities to some degree.

In a study closely related to that of Madsen and Geringer (1990) and Geringer and Madsen (1996), Johnson (1996a) investigated differential patterns of listening in musicians (band students) only, but stratified subjects on the basis of performing ability level (high or low). Subjects were asked to listen to familiar and unfamiliar band music and indicate on a CRDI device to which instrumental group they were attending at any given moment. Subjects could indicate woodwinds, brass, percussion, all, or my section. Results were consistent with previous studies. Specifically Johnson found that level of musical ability did have a significant effect on subjects’ focus of attention, and suggested that differential patterns of listening to music may be influenced by musical training.
It is evident from these studies in the perception of musical nuance that musicians’ and nonmusicians’ abilities may be regarded as similar or different depending upon the specific listening task considered. For example, musicians and nonmusicians generally tend to perceive certain global features in music (e.g., tension) similarly, and also report similar aesthetic responses to musical stimuli. On the other hand, when considering listeners’ focus of attention or their perceptions of musical nuance (e.g., tempo changes, rubato, etc.), there does seem to be marked differences based upon degree of training. This conclusion implies that musical ability level may have an effect on the link between musical perception and production. This premise will be explored in the next section where research into the effect of instruction on the perception and production of musical nuance will be discussed. The section concludes with two studies most closely related to the present investigation.

Research into the Effect of Instruction on the Perception and Production of Musical Nuance

The studies discussed above indicate that certain differential patterns of listening appear to be an artifact of musical background and experience/training. Researchers have focused their efforts in this area by either examining the relationship between musical ability and musical perception/discrimination, or studying the effect of instruction in music on the perception and production of musical nuance. The former will be addressed
first by reviewing studies that have sought to determine if higher ability in musical performance corresponds with higher ability in musical perception/discrimination skills.

Geringer (1983), Ramsey (1983), and Killian (1991) examined the relationship between musical ability (musicianship) and musical perception. Geringer examined the relationship between pitch discrimination and vocal pitch matching abilities of preschool and fourth grade students. Students were placed in high, medium, and low ability groups based upon their pitch discrimination (perception) skills. Subsequent, pitch matching tasks revealed no significant differences among ability groups, although there was a marginally significant interaction (p = .10) between age group and ability level. Although preschool students' perception and performance skills did not correlate well, scores on production tasks for fourth grade students began to correspond with perception abilities, suggesting that the link between production and perception may be an artifact of maturation or education.

Unlike Geringer (1983), Ramsey (1983) found a stronger relationship between preschool children's singing ability and age, and their perceptive abilities. Specifically, Ramsey found that 3 to 5 year old subjects with high singing ability scored significantly higher on perceptive tasks relating to melodic contour, melodic rhythm, tonal center, and melodic interval than those with lower singing abilities. Age also was an indicator of perceptive ability with older students scoring significantly higher on the perception tasks.

In a similar study with older students, Killian (1991) compared junior high school choir students' sichtsinging accuracy with their ability to perceive errors in performances. Based upon sichtsinging ability determined by a
sightsinging test, subjects were placed into one of three groups: high, medium, and low ability. Subjects then were asked to listen to a performance and indicate the errors that were present. Killian found no significant difference between sightsinging ability and error detection in the high and medium groups, but found a significant difference in the low scoring group. Subjects in this group scored significantly higher on perceptive tasks (i.e., error detection) than on performance tasks. Also of note is the relationship between groups. On sightsinging tasks, the high scoring group received a perfect score, 100%; the medium group, 73%; the low group, 30%. However on error detection tasks, the high group scored 83%, the medium group 68%, and the low group, 62%. Comparably, the high scoring group remained significantly higher than the other two groups, but the low group was not significantly different from the medium group on perceptive tasks. Although not discussed, this may indicate that participation in performance ensembles may be having a positive effect on students' perceptive abilities although their performance abilities may be lagging behind.

While these studies have focused on the effect of musical ability on the perception of musical nuance, other research has sought to explicate the relationship between training in music and musical perception. In other words, these studies have sought to determine if the amount and depth of training in music positively affects the perception of musical nuance.

Nierman (1983) examined the aural skills of students taking private music instruction to ascertain the effect of training in performance on the perception of melody, meter, rhythm, and tonal imagery. Subjects, all of whom were taking private lessons, were stratified into those with 3 or more years experience and those with less than three years experience. For his
dependent measure, Nierman used parts of Gordon’s (1965) Musical Aptitude Profile that tested melody, meter, tonal imagery, and rhythm imagery. Nierman found that there were no significant differences between groups on aural perception skills. From this, he concluded that “private music instruction does not seem to bear anything but a low to moderate relationship to music perception skill” (p. 26).

Thinking that Nierman’s (1983) study may have defined instruction in music too narrowly, Pembrook and Taylor (1986) investigated further the effect of instruction on musical perception by defining musical instruction more broadly. Pembrook and Taylor’s definition of musical instruction included private music instruction, as well as performing ensembles, music theory, piano, and music history instruction as predictor variables to determine if extensive musical study positively affects aural perception. Based upon depth of musical experience (e.g., number of years participating in ensembles, taking private lessons, taking high school music theory, etc.), subjects were placed into one of three groups: high experience, moderate experience, or low experience. Contrary to Nierman’s findings, results from Pembrook and Taylor’s study indicated that level of musical experience was a significant predictor of aural perception skills. Specifically, those with high experience scored significantly higher than those with moderate or low experience. There was no significant difference found between the low and moderate experienced groups however, indicating that it may take longer for an individual’s perceptive abilities to be significantly impacted by instruction.

In a similar line of research, Sheldon and Gregory (1997) examined the long term effects of participation in performing ensembles on the perception of tempo modulation. Using a cross sectional design, Sheldon and Gregory
used subjects who were students actively participating in band, orchestra, and
choir groups at different levels of education. Subjects ranged in grade level
from fifth grade to graduate students; music faculty members also were
considered. Subjects listened to musical excerpts and indicated the degree of
tempo increase or decrease perceived on a CRDI device. Results indicated
that level of experience has a significant impact on perceptions of tempo
modulations. For the most part, the three most experienced groups’
responses clustered together and were significantly set apart from the groups
with lesser experience. Sheldon and Gregory concluded that age and level of
experience/training have significant bearing on musical perception.

Like Sheldon and Gregory (1997), work by Johnson (1996b) also has
explicated the relationship between training in music the perception of rubato
in musical performance. Subjects (musicians/trained and nonmusicians/
untrained) were asked to evaluate the appropriateness of temporal variations
in several different performances of Mozart’s Concerto for Horn and
Orchestra No. 2, movement 1. Listeners also were asked to evaluate the
soloist’s musicianship, expression, tone quality, and tempo. It was found that
musicians and nonmusicians differed significantly in evaluations of rubato,
musicianship, expression, and tone quality, but not tempo. Stratifying the
musician group further revealed that musicians with higher levels of
experience differed significantly from musicians with lesser experience in
evaluations of rubato, musicianship, expression, and tone quality, but not
tempo. Results also indicated that rubato seems to be an extremely subtle
musical nuance that may differentiate very good performances from the very
finest. Most importantly, it was evidenced that musical training had a
significant effect on the perception of this nuance.
Although there seems to be converging evidence to indicate that instruction in music positively affects the perception of musical nuance, little research is available that examines its effect on the perception of musical expression. Two important papers, however, deserve discussion: one, a study of the effect of training on the communication of musical expression by Kendall and Carterette (1990); two, a report by the National Assessment of Educational Progress (NAEP). The former will be reviewed first, followed by a discussion of the latter.

In a series of studies, Kendall and Carterette (1990) examined the link between performer and listener to determine the impact that musical training has on the perception of musical expression. Kendall and Carterette had subjects engage in a number of perceptual tasks that involved: 1) categorizing excerpts in terms of expressivity (i.e., non-expressive, appropriate expression, overly expressive); 2) matching an instrumental version of the selection with a piano performance of the same selection based on expressive qualities; 3) categorizing synthesized performances with timing variations (i.e., no timing deviation, appropriate timing deviation, and too much timing deviation); and, 4) rating performed and synthesized excerpt’s expression on a scale from 0 (without) to 100 (great). Results from the first experiment (categorization) indicated that musicians and nonmusicians performed similarly. That is, there were no significant differences found in the ability to categorize musical performances as non-expressive, expressive, or overly expressive. The study found that discrimination between mechanical and expressive performance was relatively easy, whereas discrimination between appropriate and overly expressive performances was much more difficult. It also was found that synthesized excerpts containing variations in timing only, rather than
dynamic variations, elicited higher ratings of expression, suggesting that tempo may be a more salient feature of expression than intensity. However, expressive performances were perceived more readily when a combination of nuances varied within the performance. In other words, subjects were better able to perceive expression when performances contained variations in tempo, dynamics, and articulation.

Although Kendall and Carterette's (1990) study indicates that there is no significant difference between musicians and nonmusicians in their ability to perceive expression in music, the results of the study need to be qualified somewhat. As nonmusicians were included only in the categorization portion of the study, it is difficult to conclude that musicians and nonmusicians are entirely alike in their perception of musical expression. While it may be that nonmusicians can discriminate accurately between non-expressive and expressive performances, asking nonmusicians to engage in more discriminating perceptive tasks, such as evaluating expressive performances, may render different results. This could explain the apparent discrepancy between Johnson's (1996b) findings and those of Kendall and Carterette. Thus, differences or similarities may exist between these two groups based on the particular task being executed.

1. It is interesting to note that this finding parallels that of Gerard and Drake (1990) who found that young children are rarely able to reproduce intensity differences when asked to reproduce simple temporal structures. Again, the phenomenon seems mitigated by focus of attention in that the processing of temporal information appears to dominate attentional resources. Gerard and Drake, in discussing this, concluded that it is only when temporal information is mastered that attention can be allocated to intensity differences. Thus, even at a young age, temporal functioning appears to be a more salient feature of musical perception.
To assess the impact of music curricula on the perceptive and performance skills of students, the National Assessment of Educational Progress (NAEP) constructed an assessment measure to determine what students “know and can do” in music (National Assessment Governing Board, 1994, p. i). The test was administered to eighth grade students in 1997 as a comprehensive measure that would assesses these students’ abilities to perceive, create, and perform music. Students were selected from throughout the United States. All students were tested on their ability to perceive music regardless of their participation in music classes. However, only those students actively participating in performing were assessed on their ability to perform and create music. Results of the measure indicated that there was no significant difference in perceptive abilities between those receiving musical instruction every day and those not receiving musical instruction.

In regard to students’ ability to produce and create music, only students with performance experience either in or out of school were considered. These students were evaluated on their ability to perform, create, and improvise. Items did not require facility on an instrument in order to respond. Responses were judged by a panel of experts and scored as inadequate, limited ability demonstrated, adequate ability demonstrated, or developed ability demonstrated. In terms of students’ ability to sing with expression, results indicated that 77% were deemed as having inadequate or limited ability. This may be explained somewhat by students’ ability to sing with proper tone quality and accurate pitch/intonation where results indicated that 73% and 62%, respectively, were judged as demonstrating inadequate or limited ability. In terms of rhythmic accuracy, however, only 39% were deemed as inadequate or limited.
Taken at first glance, the NAEP results would seem to indicate that performing ensemble participation has little positive effect on students’ musical perception and production abilities. However, it is important to note that although it was not reported as significantly different, those participating in instrumental ensembles scored over one standard deviation higher\(^2\) than those who had never enrolled in an instrumental ensemble in regard to perceptive abilities (\(M = 178, \text{SD} = 2.3; M = 142, \text{SD} = 1.4\); respectively). Likewise, although not as pronounced, those participating in choral ensembles scored over one half a standard deviation higher on perceptive tasks than those who had never participated in music (\(M = 163, \text{SD} = 2.3; M = 144, \text{SD} = 1.5\); respectively). With this being considered, the effect of performing ensemble participation on students’ perceptive abilities cannot be dismissed outright.

Furthermore, the results pertaining to performing ability deserve scrutiny. First, as the NAEP assessment utilized singing to assess performance ability, it may be that instrumentalists were significantly disadvantaged. That is, because instrumentalists are not used to singing everyday in performing class, they might be less confident in their vocal abilities. Had the NAEP report assessed their musical ability on their respective instrument, findings might have been substantially different. Second, as the performance studies discussed above indicate (Palmer, 1996a, 1996b; Sloboda 1983, 1985), the ability to perform expressively takes significant

\(^2\) Scores from the NAEP assessment are reported on a scale based Item Response Theory determined to have a mean of 150, standard deviation of 35, and range of 0 to 300.
time to develop. Because most performing ensemble students would have been engaged in performing ensembles for only two to four years, it seems reasonable to speculate that these students would be just beginning to develop the fundamental skills of basic musicianship (e.g., performing with proper tone quality, performing with proper rhythm, playing in tune, reading music, etc.). It would be expected that these skills, like any other, would improve with time and experience. However, because the NAEP only assessed this particular age group, the effect of prolonged engagement in performing ensembles on the production of musical expression cannot be gleaned. Third, because the NAEP assessment included only participants in performing ensembles, how these students’ skills compare to those not participating in performing ensembles is indeterminable. Without this comparison, it cannot be concluded that each of these populations is equivalent in terms of their ability to perform music. Evidence from the report already suggests that a majority of those with performing ensemble experience are performing with rhythmic accuracy, however, this may be an artifact of maturation rather than instruction. If those not engaged in performing activities were assessed, more insight into the effect of musical training would be allowed. Furthermore, in regard to assessing the performance of musical expression, not only is it necessary to compare these two populations, it also is necessary to control for the confounding effects of tone quality and pitch/intonation. Consequently, a device that controls for these confounding influences might help in investigating this phenomenon. Furthermore, if this device could equalize, as much as possible, the technique necessary for performance so that musicians and nonmusicians could be compared, more insight into the differences between these groups could be

49
gleaned. The remainder of this section discusses work by researchers that have found a particular device to meet this end: the *Instant Pleasure* software program.

Using the *Instant Pleasure* software program, Rodriguez (1995) examined children's ability to perceive and produce musical expression. Subjects were kindergarten, second, and fourth grade students in two suburban middle class and upper middle class schools. Those studying music privately were excluded from the study. The perception phase of the study used mechanical and expressive renditions of children's songs for stimuli. Each subject was presented with three performances of the same song in a 2 + 1 oddity paradigm. That is, each subject heard two identical renditions and a different rendition, and was asked to identify the rendition that was different than the other two in terms of its expressive qualities. Results from the perception phase of the study indicated that subjects were able to discriminate between mechanical and expressive performances with 60% accuracy overall. However, this ability was shown to increase with age: kindergartners achieved just above chance levels at 46%, second grade achieved 63%, and fourth grade achieved 70%. The production phase of the study utilized *Instant Pleasure* software to allow students to engage in performing activities. Because the *Instant Pleasure* software allows for control over tempo, rhythm, and dynamics, subjects had control over the most salient features of musical expression. Subjects were allowed to choose a well known children's song that they recognized, practice performing it, and record their musical version. Each subject was instructed to perform the piece "the way [they thought] it should go" (p. 81). Each subject's final performance was rated, subsequently, by an expert for musicality/expressiveness. Results of the performance phase
of the study indicated that older subjects were better and more consistent in their use of expressive devices. These ratings were determined to be developed substantially less than perceptive abilities however. Rodríguez concluded that the underlying knowledge systems underlying perception and production abilities were not evident nor supported by the data in this study.

Although not examining the relationship between perception and production specifically, Flowers, Dregalla, and Jackson (1993) compared adult musicians and nonmusicians ability to evaluate performances and perform musically by using the Instant Pleasure software program. After an introduction to the software, subjects were asked to record a version of “Ode to Joy” to their own musical satisfaction. Recordings were placed into one of three groups based upon the subject’s musical background. One group consisted of university music faculty and staff with extensive keyboard experience, a second group consisted of music majors with minimal keyboard experience, and a third group consisted of nonmusic majors and university staff with little or no keyboard experience. A total of 106 listeners (53 graduate music students and 53 graduate or upper division education majors) evaluated the performances in terms of musicality and piano skills. In terms of evaluation, results indicated that the groups did not differ significantly in their evaluation of musicality or piano skills, although nonmusicians tended to use a smaller range than musicians in their evaluation. However, in terms of performance ability, significant differences were found between each performing group. Those with more musical experience received significantly higher ratings in both musicality and piano skills. That is,
university faculty and staff with extensive keyboard experience received the highest ratings, followed by music majors with minimal keyboard experience, followed by nonmusicians with minimal keyboard experience.

The findings of Flowers et al. (1993) would indicate that musical experience has some bearing on musical production tasks, yet may not be as prevalent in regard to the perception of musical expression. Also significant, is that performing ability using the *Instant Pleasure* software program is predicated upon keyboard experience. That is, those who had extensive training/expertise on piano scored significantly higher than those who did not have this background. This is not surprising as technique is always a concomitant of the ability to produce music well (Sloboda, 1994). It is to be expected, then, that those who have had this background would have a more expressive range based upon their familiarity with the touch and feel of the keyboard. In light of this, the present investigation controlled for this variable by excluding pianists from the performance phase of the study.

Both Rodriguez (1995) and Flowers et al. (1993) demonstrate that the *Instant Pleasure* software program can be used effectively with both musicians and nonmusicians in assessing the ability to perform with expression. By equalizing technique across these populations, differences in productive and perceptive tasks can be elucidated. It is through studying these differences that the effect of training on perceptive and productive musical tasks can be gleaned.
Summary

The study of performance practice has yielded great insight into the constituents of an expressive performance. Research has demonstrated that performers deliberately manipulate expressive devices during performance to correspond with underlying musical structures. Additionally, those with extensive background in performance have been shown to communicate musical intent more precisely by controlling the use of expressive parameters and linking these parameters to salient structural features within the musical context. It has been shown that a performer's knowledge of these structural events will assist in determining where and to what degree variations in musical nuance will be implemented. Thus, a more thorough understanding of the underlying musical structure will positively affect the performer's use of expressive devices. It has been demonstrated that the use of these expressive devices assists in the perception of music. For example, listeners have been shown to identify more accurately meter and phrase groupings when structural intent corresponds with expressive variation. However, the effect of musical production on musical perception is not quite so clear.

Unfortunately, there is little extant research available that explores the effect of performing ensemble participation on subsequent perception and production tasks. Furthermore, even less research has examined if participation in performing activities transfers positively to the production and perception of musical expression. Because much of the musical instruction in music education curricula uses performance as the primary vehicle for study, it is imperative to understand the effect that this is having on the subsequent perceptive and productive abilities of students. The
present study adds to the empirical knowledge in this field by examining the
effect of performing ensemble participation on students' ability to perceive
and produce expression in a musical performance.
CHAPTER 3

METHODS AND PROCEDURES

Purpose

The purpose of this study was to examine the effects of school performing ensemble experience (i.e., band, choir, or orchestra participation) on the ability to perceive expression in music and to perform a musical passage expressively using a minimal skill musical device. In order to assess the effect of the performing ensemble on generative and receptive abilities, the present investigation compared collegiate level subjects who had prior experience in school performing ensemble courses with subjects who had not. For the purposes of this study, musicians were defined as those subjects whose only musical training took place in schools (i.e., no private lessons, not self taught, etc.) and who had at least one year of performing ensemble experience in high school. Nonmusicians were defined as those subjects who had no previous formal training in music.

In order to address both performance and perception, the study was divided into two phases. Phase I addressed the generation of musical expression, while Phase II examined the perception of musical expression. Because expression is contingent to a large degree on the accuracy of musical
performance, an examination of musical accuracy was included in each phase of the study as well. Specific research questions that were addressed are as follows:

**Phase I:**

1. Is there a significant difference between subjects who have had school performing ensemble experience and subjects who have not in the ability to perform a musical passage with expression?

2. Is there a significant difference between subjects who have had school performing ensemble experience and subjects who have not in the ability to perform a musical passage accurately?

3. Does the number of years participating in a school performing ensemble affect the ability to perform a musical passage with expression?

4. Does the number of years participating in a school performing ensemble affect the ability to perform a musical passage accurately?

**Phase II:**

5. Is there a significant difference between subjects who have had school performing ensemble experience and subjects who have not in their ratings of perceived expression in a musical passage?

6. Is there a significant difference between subjects who have had school performing ensemble experience and subjects who have not in their ratings of perceived accuracy in a musical passage?

7. Does the number of years participating in a school performing ensemble affect ratings of perceived expression in musical passage?

8. Does the number of years participating in a school performing ensemble affect ratings of perceived accuracy in a musical passage?
Participants

The target population for this study was students who have graduated from American high schools and possess the basic musical background provided by schooling. As most school curricula treat music as an elective, some students' school musical background is extremely limited while others have had extensive musical training in the school setting. In order to achieve a representative sample of high school graduates with diverse musical experiences, participants were recruited from music appreciation courses offered by the School of Music at The Ohio State University ("Music Cultures of the World"). As these courses are open only to those not majoring in music, the enrollment consists of students with various musical backgrounds and diverse school musical experiences. Participation in Phase I and Phase II of the study was limited to subjects with no prior musical experience/training either in or out of school (nonmusicians), and subjects whose only musical training was through performing ensembles in the school setting (i.e., band, orchestra, or choir classes).

Demographics

A total of 56 subjects participated in Phase I of the experiment. Nonmusicians (N=28) consisted of Freshmen (N=11), Sophomores (N=10), Juniors (N=6), and Seniors (N=1). 15 were male, 13 female. Musicians (N=28) consisted of Freshmen (N=6), Sophomores (N=11), Juniors (N=5), and Seniors (N=6). 12 were male, 16 female. Musician subjects reported participating in either school Band (N=18) or Choir (N=10) courses; no school Orchestra
students volunteered for Phase I. Musicians reported participating in school ensembles for two (N=5), three (N=1), four (N=3), five (N=4), six (N=5), and seven (N=10) years.1

A total of 193 subjects participated in Phase II of the study, 138 of which matched the study’s operational definition of musician or nonmusician. Of the 138 remaining subjects, 9 were eliminated because of incomplete response forms or failure to take the task seriously (e.g., circling 5 for every response on the Likert scale response sheet). The remaining subjects (N=129) for Phase II were divided into nonmusicians (N=56) and musicians (N=73) based on the same criterion employed in Phase I. Nonmusicians consisted of Freshmen (N=27), Sophomores (N=12), Juniors (N=8), and Seniors (N=9). 36 were male, 20 female. Musicians consisted of Freshmen (N=19), Sophomores (N=30), Juniors (N=16), and Seniors (N=8). 34 were male, 39 female. Musician subjects reported performing in school Band (N=42), Orchestra (N=4), and Choir (N=27) courses. Participation in school ensembles was reported as one (N=11), two (N=11), three (N=5), four (N=6), five (N=5), six (N=8), seven (N=23), eight (N=3), and nine (N=1) years. The subjects with eight and nine years of participation indicated beginning school performance in the fifth and fourth grade, respectively.

1. Years of participation was calculated by adding high school and middle school years together (e.g., a subject with 1 year of high school participation and 2 years of middle school participation was considered to have 3 years of participation).
The Experimental Apparatus

Phase I of the study examined research questions pertaining to the performance of musical passages. As instrumental technique and skill confound the investigation of performance activities, it was necessary to employ a performing apparatus that might equalize performing technique between musicians and nonmusicians. The apparatus used to meet this end was the *Instant Pleasure* software program which was connected to a MIDI keyboard. This software was designed to allow nonmusicians and musicians to perform with minimal skill requirements. Using the software, the performer can depress any key or keys on a MIDI keyboard to trigger the onset of each note or chord in a song file. The software, thus, "performs" the correct melody and accompaniment regardless of what key is depressed on the MIDI keyboard. By giving the performer control over note and chordal onsets, the software allows the performer to manipulate the rhythm and tempo of any given performance. The performer also has direct control over dynamic variation as the software is responsive to touch sensitive MIDI keyboards. Thus, the *Instant Pleasure* software apparatus allows the performer control over expressive devices such as tempo and dynamics while providing a note perfect rendering of the song.

Selection of the Musical Excerpts

The musical excerpts that each participant was asked to perform using the *Instant Pleasure* software consisted of two familiar pieces and one unfamiliar piece. In order to select the two familiar excerpts, a repertoire list
was formulated and presented to the students enrolled in the "Introduction to Music" course at the Ohio State University. As these students have similar musical backgrounds, it was assumed that their responses would parallel those of students enrolled in "Music Cultures of the World." Students were instructed to circle the songs on the repertoire list with which they were most familiar and could sing or hum from memory. The responses were reviewed, and only the selections that received 100% agreement among the participants were retained. The songs that received 100% consensus were as follows: Amazing Grace; America the Beautiful; The First Noel; Joy to the World; Jingle Bells; My Country 'Tis of Thee; Silent Night; and, Twinkle, Twinkle, Little Star. A list of these songs was presented to five music performance and five music education faculty at The Ohio State University in order to determine which of the excerpts had the greatest potential for expressive performance. Faculty were asked to rank order the excerpts in terms of expressive potential, and the top two selections with the greatest potential for expressive interpretation were retained for the experiment. The two songs selected by the judges were Amazing Grace and America the Beautiful.

The song that was selected as the unknown excerpt was the first sixteen measures of Edward MacDowell's To a Wild Rose from Woodland Sketches. The song consists of melody with chordal accompaniment in simple duple time that provides ample opportunity for expressive variation. Furthermore, the song provided a challenge to subjects in that the first phrase is slightly different in rhythm than the second. It was expected that this rhythmic variation would allow for a more accurate assessment of subjects' musical skill.
Appendix A provides full scores for the musical excerpts. The composite rhythm that is indicated beneath each score represents the rhythm that each subject needed to perform in order to render an accurate performance in terms of both melody and harmonic accompaniment. It is evident when viewing the scores that the two familiar excerpts represent two different degrees of difficulty because of the rhythmic complexity generated by the passing tones. The passing tones in *America the Beautiful* are symmetrical in nature and occur predominantly on the beat. Consequently, if the subject was able to keep a steady metronomic pulse through these passing tone measures, the rendering would be accurate. On the other hand, the passing tones in *Amazing Grace* were not as intuitive or as easy to render. In this case, subjects needed to remember where in the song file the passing tones occurred and which beats constituted the passing tones in order to trigger a rhythmically correct performance. In other words, subjects could not simply tap the metronomic pulse and achieve an accurate performance.

All performances were roughly the same length. *Amazing Grace* lasted 29 seconds, *America the Beautiful* lasted 39 seconds, and *To a Wild Rose* lasted 30 seconds.

**Method: Phase I**

Each section of the “Music Cultures of the World” course was visited by the experimenter in order to recruit subjects for Phase I. Subjects volunteered for Phase I of the experiment by signing up to participate. The experimenter contacted each subject to inquire about their musical background, and to arrange a convenient time for the experiment to take
place. Only those subjects with no previous instruction/training in music and those with a minimum of one year of high school performing ensemble experience were asked to participate.

The location of the experiment was a large classroom in the School of Music converted into a MIDI lab for the express purposes of this study. The lab consisted of four MIDI stations, each containing a set of headphones and a touch sensitive MIDI keyboard connected to a Macintosh computer running the *Instant Pleasure* software. Each MIDI keyboard was set to a piano timbre throughout the study. The lab was located away from distractions and arranged so that participants could not interfere or distract each other during the experiment.

Subjects arrived and were asked to complete a survey instrument on their musical background (Appendix B). This instrument was designed to obtain information concerning musical experience both in and out of the school setting. After completing the survey instrument, subjects were asked to read the following instructions:

In this experiment, you will be asked to perform two familiar songs and one unfamiliar song using computer software that will assist you in playing the piano. For each song you will be asked to perform in two ways: unexpressively and expressively.

For the unexpressive version of your song, try to make the performance sound as "dry" or without expression as possible. For the expressive version, try to make the performance sound as expressive or musical as possible.

You will have a maximum of 10 minutes on each song in which to practice and record both performances (i.e., the unexpressive and expressive versions). During this time you may listen to the song as much as you like, practice the song as much as you like, and record the song until you achieve a final recorded version with which you are satisfied. When you have achieved a recording with which you are satisfied, raise your hand and the experimenter will save your work. You may then move on to your next example.
Before beginning your practice work on each song you are asked to listen to the song two times in order to familiarize yourself with the song. Even if you think you know how the song goes, you are asked to do this as the computer version of the song may be different from the way you remember the song.

After you have completed the last song, please fill out the Follow Up Survey provided. Please record your assigned participant number in the space provided on the top of the survey instrument.

For your convenience a summary of the software's commands are printed below. To activate these functions simply use the mouse to click on the corresponding icon:

LISTEN.- This will play the song in its entirety. Remember this is just an example. You do not have to perform the song in this manner. You may change the speed and volume at any time during your performance to make the piece sound more or less expressive.

PERFORM.- This will allow you to practice the song by playing the keyboard.

RECORD.- This will allow you to record your rendition of the song.

STOP.- This stops the listening, performing, and recording functions and returns you to the beginning of the piece.

After subjects had read the instructions, the experimenter provided a brief training session on how to use the Instant Pleasure software. The experimenter used Twinkle, Twinkle, Little Star to demonstrate the experimental apparatus. The script for the training session was as follows:

For this experiment you will be asked to perform three different songs in two ways: expressively and unexpressively. The songs you will perform are Amazing Grace, America the Beautiful, and a song you will not know. So you will perform Amazing Grace expressively and Amazing Grace unexpressively, America the Beautiful expressively and America the Beautiful unexpressively, and the unknown song expressively and unexpressively. At the conclusion of the experiment you will have a total of six performances. The order that you will record your songs is listed on the sheet of paper I provided you. As an example, I will demonstrate a recording session using Twinkle, Twinkle, Little Star.

The first screen that you will see will look like this (Twinkle, Twinkle, Little Star's interface will be opened here). As you can see, the screen contains five boxes: Stop, Pause, Listen, Perform, and Record. These buttons function exactly as a cassette player or CD player would. The first thing you are asked to do is to listen to the song a minimum of two times in order to familiarize yourself with it. So, you will use the mouse and click on “Listen” (click on
Listen and Listen to *Twinkle* all the way through. It’s important to realize that these renditions that you will listen to are not intended to represent either expressive or unexpressive performance, but are simply there to remind you of how the song goes.

After you have listened to the song a minimum of two times, you may practice performing the song as much as you like. Use the mouse to click on “Perform” in order to practice. (Click on Perform). Now, the nice thing about this software is that any key that you press on the piano keyboard is the “right” one. (Demonstrate playing *Twinkle* hitting random keys on the MIDI keyboard). So, to play the song you might want to use one key with one finger, (demonstrate), two keys using different hands (demonstrate), or two keys using the same hand (demonstrate). What ever method is most comfortable for you is okay. The important thing is that you are able to render the song as well as you can.

The other nice thing about this program is that you can vary the speed of the song by pressing the keys faster or slower and the computer will keep up or wait for you (demonstrate different tempos). You also can change the volume by pressing the keys harder or softer (demonstrate loud and soft). When you are asked to perform expressively you might vary these parameters to make the music sound more expressive (demonstrate an expressive version of *Twinkle*). When you are asked to perform without expression you may want to eliminate these parameters to make the performance sound more “dry” (demonstrate unexpressive *Twinkle*). Why don’t you try to see if you know what I mean? (Subject was allowed to experiment with *Twinkle* until he/she felt comfortable using the device, and any questions were answered).

After you have practiced the song and feel comfortable performing it, you can record your version by clicking “Record” (Record was clicked and some of *Twinkle* performed). If you make a mistake in the performance you may click on “Stop” and start recording again at the beginning. After you have recorded a version you like, raise your hand and I will save your work.

You have a maximum of ten minutes in which to record both versions of the song; the expressive and unexpressive versions. Within this time you may listen to, practice, and record the song as many times as you would like. Are there any questions?

Subjects were then assigned to a MIDI station and allowed to adjust the volume in their headphone until it was to their liking. To control for any order effects, the order of the songs was randomized using a Latin Square design. Subjects were given the order in which they were to perform the excerpts before beginning the experiment. A ten minute time limit was imposed in which subjects were allowed to listen to, practice, and record the
expressive and unexpressive versions of each song (i.e., 10 minutes for both versions of Amazing Grace, 10 minutes for both versions of America the Beautiful, and 10 minutes for both versions of the unknown song). No subject exceeded the ten minute time limit imposed. Prerecorded listening examples were provided in order to refamiliarize subjects with the songs, but these renditions were quantized rhythmically and dynamic variation was minimal. Thus these performances were considered accurate but not musically expressive.

At the conclusion of the experiment, subjects completed a brief follow-up survey (Appendix C). On this survey, none of the subjects indicated recognizing the unknown song. Furthermore, all subjects indicated that their recordings were representative of their intentions.

Each of the final six renditions from each participant (three selections by two renditions, expressive and unexpressive) was saved as a MIDI file for a total of 336 performances. Audio cassette tapes randomly ordering the participants’ final renditions was produced for evaluation. In order to make the task less time consuming for evaluators, a random sample of 67 examples (20%) was evaluated by all judges so that inter-rater reliability could be calculated. The remaining examples were divided randomly among evaluators. Additionally, 12 examples were repeated within each judge’s tape in order to calculate the internal consistency of ratings.

Four expert musicians were recruited to evaluate the participants’ performances. Three of the adjudicators were Professors of Music Education at OSU regional campuses; one adjudicator was an Associate Professor of Music Education at Virginia Commonwealth University. Each judge received two cassette tapes that contained a random sample of the final
performances, and a reference cassette tape that provided the computer generated models that subjects used to familiarize themselves with each song. Judges were instructed to listen to the reference cassette before beginning their evaluation. Performances were rated on both musical accuracy and the overall expressiveness of each performance on a Likert scale ranging from 1 (poor) to 7 (excellent) and 1 (unexpressive) to 7 (very expressive), respectively. Judges were told that the performances were by non-music majors using a MIDI device for performance assistance. Specific instructions were as follows:

The following musical examples were performed with the assistance of a MIDI device by undergraduate students not majoring in music at the Ohio State University. This MIDI device allowed each student the opportunity to control, for the most part, the expressive qualities of each performance.

You are asked to listen to the performances and rate the overall musical accuracy and expressiveness/musicality of each interpretation. For each selection, rate musical accuracy on a scale from 1 (poor) to 7 (excellent), and the overall expressiveness/musicality on a scale from 1 (unexpressive/without expression) to 7 (highly expressive/musical). Please listen to the entire selection before making your judgment.

Before you begin, listen to the example tape provided. The examples were provided for the performers as well, and were intended only to provide a reference or “road map” for musical accuracy. In other words, the performers were allowed to deviate temporally (e.g., by using rubato, taking a different tempo, etc.) and dynamically from the provided examples. The examples contain no expressive intent.

The songs you will evaluate are Amazing Grace, America the Beautiful, and To a Wild Rose.

Each judge evaluated a total of 148 examples that were broken down into four sessions of 37 examples each. Each session lasted approximately 25 minutes. The 67 examples that were common to all judges were randomized and counterbalanced across judges to control for any order effects. In the final data analysis, the 67 examples that each judge had in common were averaged
in order to render a single mean score. All other examples were evaluated by one judge and therefore do not represent a composite rating. Inter-rater reliability and internal consistency coefficients are reported in Chapter 4.

**Method: Phase II**

Phase II used "Music Cultures of the World" courses in order to examine the research questions pertaining to the perception of musical expression. Access to these courses was granted by the instructor, and the experimenter set up a date that would be convenient to administer the experiment. A total of 12 classes were used for Phase II.

A stimulus cassette tape was constructed from the final recorded performances generated in Phase I of the study. The tape contained 54 randomly selected examples, 10 of which repeated in order to calculate internal consistency among the subjects. Examples were randomized and counterbalanced for each class in an effort to control for any order effects. The tape lasted approximately 40 minutes. Each stimulus tape was played on stereo equipment provided in School of Music classrooms. Volume levels were adjusted to an appropriate level for each classroom.

As many of the classes met at the same time, graduate students and music education faculty were recruited to assist in proctoring the experiment. Each proctor was trained in the administration of the experiment and was provided a script to follow. On the day of the experiment, proctors arrived and notified subjects that the class period was to be used for the experiment. Subjects were told that the experiment was completely voluntary and that they could choose not to participate without any prejudice to them.
whatever. A minority of the subjects elected not to participate and left class for the entire period. The subjects who elected to participate were given written instructions for the experiment which were as follows:

You are about to hear several different performances of three pieces: Amazing Grace, America the Beautiful, and a song you may not know entitled, To a Wild Rose.

You will be asked to listen to each performance and then evaluate the overall accuracy and expressiveness/musicality of the performer’s interpretation.

For each selection, rate musical accuracy on a scale from 1 (poor) to 7 (excellent), and the overall expressiveness on a scale from 1 (unexpressive/without expression) to 7 (highly expressive/musical). Please listen to the entire selection before making your judgment.

Participation in this experiment is completely voluntary and you may choose not to participate or elect to discontinue participation at any time during the experiment without prejudice.

Before we begin, an example of each song will be provided. Remember, it is possible for a performance to be faster or slower than these examples and still be considered musically accurate. The examples are intended only to provide a reference or “road map” for musical accuracy. The examples have no expressive intent.

After the subjects had read the instructions, proctors read a summary of the instructions to the subjects in case there was any confusion. Subjects were told that they were being asked to evaluate performances of Amazing Grace, America the Beautiful, and a song they might not know entitled, To a Wild Rose. Each song was to be evaluated on musical accuracy and expressiveness (i.e., how expressive or musical they thought the performance to be). For musical accuracy, 1=poor and 7=excellent; for expressiveness, 1=without expression/unexpressive and 7=highly expressive. Time was then taken to answer any questions.

In order to familiarize students with the songs, an example tape was played first that contained the computer renditions of the songs to which subjects in Phase I listened in order to familiarize themselves with the
excrites. Subjects were reminded that these examples were merely a road map or guide to musical accuracy and contained no expressive intent. After playing the example tape, the stimulus tape was played. At the conclusion of the stimulus tape, students were asked to complete a musical background survey (Appendix B) that solicited students involvement with music both in and out of the school environment.

A total of 198 subjects participated in Phase II of the experiment. It was decided to eliminate those who had training outside of the school environment (e.g., private lessons, self taught, church choir, etc.) in order to have a sample that was identical to Phase I of the study. Because of this, 55 subjects were excluded from the final analysis, leaving a total of 138. Upon examining the remaining 138 response forms, 9 response forms were determined unusable for analysis because they were incomplete or did not take the task seriously (e.g., only circling 5 throughout the experiment). After these eliminations, a total sample size of 129 was used for analysis, 56 of which were nonmusicians, 73 musicians.

**Analysis of Data**

Data from Phase I and Phase II of the study were entered into SPSS for statistical analysis. In order to answer the research questions pertaining to the production of musical expressiveness, Phase I data were analyzed independently of Phase II. Mean scores and standard deviations for the musician and nonmusician groups are presented in Chapter 4. These data were analyzed in terms of musical accuracy and musical expression. Additionally, the musician group was stratified into years of participation in
order to investigate the effect of prolonged engagement in performing ensembles. Data are presented which examine this group in terms of years of participation in a performing ensemble. Inferential statistics were calculated for each of these analyses using a Repeated Measures ANOVA method as outlined in Kennedy and Bush (1985).

In an effort to answer the research questions pertaining to the perception of musical expression, Phase II data were analyzed independent from Phase I. Mean scores and standard deviations for the musician and nonmusician group are presented in Chapter 4. As in Phase I, these data were analyzed in terms of musical accuracy and expression. Again, the musician group was stratified into years of participation in order to ascertain the effect of prolonged engagement in performing ensembles on the ability to perceive musical expression and accuracy. Inferential statistics were calculated through Repeated Measures ANOVA techniques.

In addition to the descriptive and inferential statistics provided in Chapter 4, inter-rater reliability and internal consistency coefficients are presented and discussed.
CHAPTER 4

RESULTS

The purpose of this study was to examine the effects of school performing ensemble experience (i.e., band, choir, or orchestra participation) on the ability to perceive expression in music and to perform a musical passage expressively using a minimal skill musical device. In order to assess the effect of the performing ensemble on generative and receptive abilities, the present investigation compared collegiate level subjects who had prior experience in school performing ensemble courses with subjects who had not. In order to address both performance and perception, the study was divided into two phases. Phase I addressed the generation of musical expression, while Phase II examined the perception of musical expression. Because expression is contingent to a large degree on the accuracy of musical performance, it was decided to include musical accuracy as a dependent variable in each phase of the study as well.

In Phase I, an experiment was devised in which subjects were asked to perform Amazing Grace, America the Beautiful, and an unknown song, To A Wild Rose, under two expressive conditions: expressively and unexpressively. Subjects were divided for analysis into a musician group (N=28) and a nonmusician group (N=28) based upon their musical background. Musicians included subjects whose only musical experience was
in school performing ensembles (a minimum of 1 year of high school), while nonmusicians included subjects who had no prior training in music. The 336 total performances generated (56 subjects X 6 performances each) were divided among four expert musicians for adjudication purposes. 20% of the performances were held in common between judges in order to calculate inter-rater reliability coefficients, while 12 performances repeated within each judge in order to calculate internal consistency. Judges rated performances on a scale of 1 to 7 for musical expression and musical accuracy (Appendix D).

In Phase II, 54 performances from Phase I were randomly selected for use as musical stimuli in an experiment devised to examine the perception of musical expression and accuracy. Subjects, musicians (N=73) and nonmusicians (N=56), listened to the performances and rated each on a scale of 1 to 7 for musical expression and musical accuracy. As in Phase I, 10 of the examples repeated so that internal consistency could be calculated.

The information in this chapter is divided into four major sections. The first section presents the effect of order, gender, and type of ensemble participation (i.e., band, orchestra, or choir) found in both phases of the study. The second section presents inter-rater and internal consistency calculations from Phase I and Phase II. The third section presents data from Phase I of the study. Each research question from this phase is approached individually with means and standard deviations presented first, followed by the corresponding inferential statistics. Finally, the fourth section presents the data generated from Phase II of the study. Each research question from this phase is approached individually with means and standard deviations presented first, followed by the corresponding inferential statistics.
The Effect of Order, Gender, and Type of Ensemble Participation

In an effort to make subsequent analyses more parsimonious, the effects of order, gender, and type of ensemble participation (i.e., band, orchestra, or choir) were examined to determine if any significant variance in the data was attributable to these variables.

The effects of order in both Phase I and Phase II were examined in relation to expression and accuracy ratings. A Repeated Measures Analysis of Variance procedure was used to determine if any significant differences could be attributed to order effects. For Phase I, effects of order were found to be non-significant for both expression ($F(3, 52) = 0.88, p = .46$) and accuracy ($F(3, 52) = 0.34, p = .80$). For Phase II, in examining rating of expression and accuracy, again, order was found to be non-significant ($F(3, 125) = 2.72, p = .36$ and $F(3, 125) = 0.69, p = .50$, respectively). Based on these results, it was concluded that there were no significant differences attributable to order in the study.

A Repeated Measures Analysis of Variance for gender was conducted to ascertain any differences in expression and accuracy scores that could be accounted for by gender. Phase I expression ratings revealed a non-significant effect for gender ($F(1, 54) = 0.02, p = .89$), as did accuracy ratings ($F(1, 54) = 0.01, p = .98$). Likewise, Phase II analyses based on gender revealed no significant effects (expression ratings, $F(1, 127) = 0.18, p = .68$; accuracy ratings, $F(1, 127) = 0.06, p = .80$).

Finally, type of group participation (i.e., band, orchestra, or choir) was examined using the same Repeated Measures Analysis of Variance technique. Non-significant differences were revealed for both expression ratings ($F(1, 26)$
= 0.01, p = .93) and accuracy ratings ($F(1, 26) = 0.01, p = 0.93$) in Phase I. Likewise, in Phase II, there were no significant differences on ratings of expression ($F(2, 70) = 0.96, p = .39$) or accuracy ($F(2, 70) = 0.03, p = .96$) based on type of ensemble participation.

Based on the non-significant nature of these findings, subsequent analyses of Phase I and Phase II data eliminated these variable from the statistical model. Thus, analyses were able to focus on the main variable of group membership (i.e., musician or nonmusician) and the number of years of participation in school performing ensembles.

**Inter-rater Reliability and Internal Consistency**

For both Phase I and Phase II, inter-rater reliability and internal consistency were calculated using Spearman’s rho correlational procedures. The Spearman’s rho statistical procedure was chosen over the Pearson Product-Moment correlation coefficient because of the ordinal nature of the data, and because it was desired to know “the degree to which subjects maintain[ed] the same relative position” (McMillan & Schumaker, 1993) both across and within judges for Phase I and within subjects for Phase II.

A moderate relationship was evidenced between the judges on ratings of expression ($r_{s(avg)} = .57$) and ratings of accuracy ($r_{s(avg)} = .69$) in Phase I of the study. However, all judges proved to be highly internally consistent on both ratings of expression ($r_s = .84$) and accuracy ($r_s = .88$).
Subjects in Phase II were found to have a much lower degree of internal consistency than the judges in Phase I on ratings of expression ($r_s = .35$) and accuracy ($r_s = .36$). Because of this low correlation, it was decided to stratify the subjects based on group membership (i.e., musicians and nonmusicians) to see if any differences existed in terms of internal consistency. Stratifying the groups revealed that musicians were more internally consistent than nonmusicians in terms of ratings of expression ($r_s = .39$ to $r_s = .29$, respectively) and ratings of accuracy ($r_s = .44$ to $r_s = .20$, respectively).

It is important to note that these low levels of internal consistency and inter-rater reliability are consistent with the extant research in performance adjudication (Abeles, 1973; Burnsed, Hinkle, King, 1985; Fiske, 1975, 1977a; Wapnick, Flowers, Alegant, & Jasinskas, 1993). Fiske (1975, 1977a), for example, has noted that high levels of inter-rater reliability and internal consistency are often difficult to obtain when judging music performance, even when professional musician adjudicators are employed. In a study investigating the consistency of judges' evaluations of trumpet performances, Fiske (1975) found that professional musicians exhibited an internal consistency of .25 on test-retest items. Inter-rater reliability also was low, averaging .50. Fiske later concluded that individual "adjudicators rarely can demonstrate consistency on more than twenty-five percent of rated performances" and that a panel of judges "could achieve an average... consistency of approximately sixty to seventy percent" (1977b, p. 24).
In light of this, adjudicators for Phase I demonstrated remarkable internal consistency and reasonable inter-rater reliability. It is important to note that the adjudicators were not trained in the evaluation of the musical stimuli, however. Had they been trained, it would be reasonable to expect that reliability would have increased.

Subjects in Phase II are more similar to the findings presented by Fiske (1975, 1977a, 1977b) in regard to internal consistency. Although low, the combined correlation is still higher than that reflected in the related literature. When stratifying the group, it is interesting to note that only the nonmusician listeners' ratings of musical accuracy fell below the 25% consistency mark. (See Chapter 5 for a more thorough discussion of these findings.)
Presentation of the Data from Phase 1

Results pertaining to research questions one through four are presented below. These data are intended to elucidate the nature of the school performing ensemble experience as it relates to the ability to perform a musical passage expressively and accurately. For clarity, it was decided to treat the ability to perform with expression and the ability to perform accurately as separate analyses.

Research Question 1: Is there a significant difference between subjects who have had school performing ensemble experience and subjects who have not in the ability to perform a musical passage with expression?

Means and standard deviations of subjects' ability to perform expressively are presented in Table 1. Examination of this table reveals that the most expressive performances were evidenced by the musician group on America the Beautiful ($M=5.31$) and Amazing Grace ($M=4.55$). More importantly, when asked to perform a song expressively, the musician group was able to perform with more expression than the nonmusician group on each song (Amazing Grace, musicians $M=4.55$, nonmusicians $M=3.54$; America the Beautiful, musicians=5.31, nonmusicians, $M=4.42$; unknown song, musicians $M=4.37$, nonmusicians, $M=3.49$). In regard to unexpressive performances, musicians also were able to perform with less expression than nonmusicians on Amazing Grace (musicians, $M=2.89$; nonmusicians, $M=3.01$) and America the Beautiful (musicians, $M=3.45$; nonmusicians $M=3.53$).
<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>Amazing Grace</th>
<th></th>
<th>America the Beautiful</th>
<th></th>
<th>Unknown</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Expressive</td>
<td>Unexpressive</td>
<td>Expressive</td>
<td>Unexpressive</td>
<td>Expressive</td>
<td>Unexpressive</td>
</tr>
<tr>
<td>Musicians</td>
<td>28</td>
<td>4.55</td>
<td>2.89</td>
<td>5.31</td>
<td>3.45</td>
<td>4.37</td>
<td>3.32</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.70</td>
<td>1.64</td>
<td>1.45</td>
<td>2.13</td>
<td>1.33</td>
<td>1.93</td>
</tr>
<tr>
<td>Nonmusicians</td>
<td>28</td>
<td>3.54</td>
<td>3.01</td>
<td>4.42</td>
<td>3.53</td>
<td>3.49</td>
<td>2.88</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.67</td>
<td>1.36</td>
<td>1.91</td>
<td>1.97</td>
<td>1.93</td>
<td>1.70</td>
</tr>
<tr>
<td>Total</td>
<td>56</td>
<td>4.04</td>
<td>2.95</td>
<td>4.87</td>
<td>3.49</td>
<td>3.93</td>
<td>3.10</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.75</td>
<td>1.49</td>
<td>1.73</td>
<td>2.03</td>
<td>1.70</td>
<td>1.81</td>
</tr>
</tbody>
</table>

Table 1: Means (M) and Standard Deviations (SD) of Expression Scores for Musician and Nonmusician Performers Reported by Song and Expressive Condition.
However, nonmusicians ($M=2.88$) performed with less expression on the unknown song than did musicians ($M=3.32$), when asked to perform without expression.

It is evident from Table 1 that all subjects were able to differentiate expressive performances from unexpressive performances and communicate more or less expression when asked to do so. However, musicians were able to make more of a differentiation between expressive and unexpressive performances than nonmusicians on each song: *Amazing Grace* (musicians, $\Delta=1.66$, nonmusicians, $\Delta=0.53$); *America the Beautiful* (musicians, $\Delta=1.86$, nonmusicians, $\Delta=0.89$); and the unknown (musicians, $\Delta=1.05$; nonmusicians, $\Delta=0.61$). To explicate this relationship further, Table 2 presents means and standard deviations for musicians’ and nonmusicians’ expressive and unexpressive performances collapsed across songs. Similar findings to Table 1 are evidenced: musicians, $\Delta=1.52$, nonmusicians, $\Delta=0.68$.

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>Expressive</th>
<th>Unexpressive</th>
</tr>
</thead>
<tbody>
<tr>
<td>Musicians</td>
<td>28</td>
<td>4.74</td>
<td>3.22</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.41</td>
<td>0.23</td>
</tr>
<tr>
<td>Nonmusicians</td>
<td>28</td>
<td>3.82</td>
<td>3.14</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.47</td>
<td>0.28</td>
</tr>
</tbody>
</table>

| Total         | 56 | 4.29       | 3.18         |
|               |    | 0.42       | 0.23         |

Table 2: Means ($M$) and Standard Deviations ($SD$) of Expression Scores for Musician and Nonmusician Performers Collapsed Across Songs.

79
Expression ratings were subjected to a three way (2 x 3 x 2), Type III Repeated Measures Analysis of Variance to investigate differences. The results, presented in Table 3, indicated four sources of statistical significance: (a) main effects for group, $F(1, 54) = 4.46, p < .05$; (b) main effects for song, $F(2, 108) = 6.37, p < .01$; (c) main effects for expressive condition, $F(1, 54) = 28.47, p < .001$; and (d) first order interaction effects between group membership and expressive condition, $F(1, 54) = 4.23, p < .05$.

Figure 1 depicts the first order interaction between group membership and expressive condition collapsed across songs. To explicate further the nature of the differences between the four means, Tukey’s HSD post hoc follow up procedures were employed in order to make a substantial number of pairwise comparisons while holding alpha at an experimentwise level of .05 (described in Kennedy and Bush, 1985). When comparing musicians’ expressive performances with their respective unexpressive performances, significant differences were evidenced (Tukey $a = 3.94, p < .05, df = 54$). Additionally, a difference approaching statistical significance was evidenced between musicians expressive performances and nonmusicians expressive performances (Tukey $a = 2.41, p < .10, df = 54$). These two findings suggest that musicians not only were able to differentiate significantly between expressive and unexpressive performances, but also were able to communicate significantly more expression than nonmusicians in expressive performances. Other than these two differences, no other significant differences were found between the means.

In regard to the significant main effect attributable to song, Figure 2 presents the overall mean scores for songs collapsed across groups and expressive condition. In subsequent post hoc procedures, it was revealed that
<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>p</th>
<th>Eta Sqr.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Between Subjects</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Group (G)</td>
<td>1</td>
<td>21.3778</td>
<td>4.46</td>
<td>.039</td>
<td>0.08</td>
</tr>
<tr>
<td>Subjects (S) / G</td>
<td>54</td>
<td>4.7978</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Within Subjects</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Song (SNG)</td>
<td>2</td>
<td>16.7500</td>
<td>6.37</td>
<td>.002</td>
<td>0.11</td>
</tr>
<tr>
<td>SNG x G</td>
<td>2</td>
<td>0.498</td>
<td>0.19</td>
<td>.828</td>
<td>0.00</td>
</tr>
<tr>
<td>S x SNG / G</td>
<td>108</td>
<td>2.6315</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Expression (EXP)</td>
<td>1</td>
<td>101.5859</td>
<td>28.47</td>
<td>.000</td>
<td>0.35</td>
</tr>
<tr>
<td>EXP x G</td>
<td>1</td>
<td>15.1093</td>
<td>4.23</td>
<td>.044</td>
<td>0.07</td>
</tr>
<tr>
<td>S x EXP / G</td>
<td>54</td>
<td>3.5684</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SNG x EXP</td>
<td>2</td>
<td>2.1465</td>
<td>0.93</td>
<td>.397</td>
<td>0.02</td>
</tr>
<tr>
<td>G x SNG x EXP</td>
<td>2</td>
<td>0.9317</td>
<td>0.41</td>
<td>.668</td>
<td>0.01</td>
</tr>
<tr>
<td>S x SNG x EXP / G</td>
<td>108</td>
<td>2.3011</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Total** 335

Table 3: 2 x 3 x 2 Repeated Measures Analysis of Variance of Expression Scores by Group (Musicians vs. Nonmusicians).
Figure 1: Mean Expression Ratings for Musicians and Nonmusicians Collapsed Across Songs.
Figure 2: Mean Expression Ratings for Songs Collapsed Across Group and Expressive Condition.
performances of *America the Beautiful* received significantly higher ratings of expression overall than *Amazing Grace* (p < .01) and the unknown song (p < .01).

Also, a significant main effect is evidenced for group (p < .05), indicating that musicians were rated significantly higher than nonmusicians when collapsing across expressive conditions and song. However, when viewing findings related to the performance of expression, the main effect attributable to group is not as meaningful as the findings related to the first order interaction between group and expressive condition, as the main effect for group compares the mean of expressive and unexpressive performances combined.

**Research Question 2:** Is there a significant difference between subjects who have had school performing ensemble experience and subjects who have not in the ability to perform a musical passage accurately?

Table 4 presents mean and standard deviation scores for subjects’ ability to perform a musical passage accurately. Highest accuracy ratings were evidenced for musicians performing *America the Beautiful* expressively (M=5.63), followed by a tie between musicians performing the unknown song expressively (M=5.09) and nonmusicians performing *America the Beautiful* expressively (M=5.09). The lowest accuracy ratings were for musicians performing *Amazing Grace* unexpressively (M=3.39). With the exception of this unexpressive version of *Amazing Grace* (musicians M=3.39; nonmusicians M=3.76), musicians were deemed consistently more accurate in their performances than nonmusicians across all songs and expressive conditions. Additionally, both groups tended to perform *America the*
<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>Amazing Grace</th>
<th></th>
<th>America the Beautiful</th>
<th></th>
<th>Unknown</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Expressive</td>
<td>Unexpressive</td>
<td>Expressive</td>
<td>Unexpressive</td>
<td>Expressive</td>
<td>Unexpressive</td>
</tr>
<tr>
<td>Musicians</td>
<td>28</td>
<td>4.73</td>
<td>3.39</td>
<td>5.63</td>
<td>4.95</td>
<td>5.09</td>
<td>4.23</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.32</td>
<td>1.53</td>
<td>1.24</td>
<td>2.08</td>
<td>1.54</td>
<td>2.11</td>
</tr>
<tr>
<td>Nonmusicians</td>
<td>28</td>
<td>3.86</td>
<td>3.76</td>
<td>5.09</td>
<td>4.45</td>
<td>4.31</td>
<td>3.64</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.07</td>
<td>1.64</td>
<td>1.77</td>
<td>2.04</td>
<td>1.93</td>
<td>1.70</td>
</tr>
<tr>
<td>Total</td>
<td>56</td>
<td>4.29</td>
<td>3.58</td>
<td>5.36</td>
<td>4.65</td>
<td>4.70</td>
<td>3.94</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.27</td>
<td>1.58</td>
<td>1.54</td>
<td>2.05</td>
<td>1.77</td>
<td>1.93</td>
</tr>
</tbody>
</table>

Table 4: Means (M) and Standard Deviations (SD) of Accuracy Scores for Musician and Nonmusician Performers Reported by Song and Expressive Condition.
Beautiful (collapsed M=5.00) more accurately than the unknown song (collapsed M=4.32) and Amazing Grace (collapsed M=3.94). Interestingly, a trend of performing the unexpressive version of each song less accurately than the expressive version is evidenced in each group across all songs.

Accuracy ratings were subjected to a three way (2 x 3 x 2), Type III Repeated Measures Analysis of Variance to investigate differences. The results, presented in Table 5, indicated two sources of statistical significance: (a) main effects for song, F(2, 108) = 13.90, p < .001; and (b) main effects for expression F(1, 54) = 13.99, p < .001. It is important to note that the difference between groups on the ability to perform accurately approached statistical significance, F(1, 54) = 2.84, p < .10.

Thus, in examining the main effect for song by collapsing across groups and expressive condition indicated that subjects were able to perform America the Beautiful (collapsed M=5.00) significantly more accurately (p < .05) than Amazing Grace (collapsed M=3.94) and the unknown song (collapsed M=4.32). Also, subjects performed expressive versions of songs more accurately than unexpressive versions (collapsed M=4.78 and collapsed M=4.05, respectively, p < .01).

In regard to differences between groups, a main effect that approached statistical significance (p < .10) was evidenced between musicians (overall M=4.65) and nonmusicians (overall M=4.18). This suggests that musicians were slightly more accurate in their performances overall. It is important to note that if musicians performed the unexpressive version of Amazing Grace more accurately, these differences would be more pronounced. Nonetheless, a specific trend can be evidenced when collapsing across expressive condition and plotting groups by songs (see Figure 3). Although none of the differences
<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>p</th>
<th>Eta Sgr.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Between Subjects</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Group (G)</td>
<td>1</td>
<td>18.4573</td>
<td>2.84</td>
<td>.098</td>
<td>0.05</td>
</tr>
<tr>
<td>Subjects (S) / G</td>
<td>54</td>
<td>6.5010</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Within Subjects</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Song (SNG)</td>
<td>2</td>
<td>32.7111</td>
<td>13.90</td>
<td>.000</td>
<td>0.21</td>
</tr>
<tr>
<td>SNG x G</td>
<td>2</td>
<td>1.2864</td>
<td>0.55</td>
<td>.581</td>
<td>0.01</td>
</tr>
<tr>
<td>S x SNG / G</td>
<td>108</td>
<td>2.3544</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Expression (EXP)</td>
<td>1</td>
<td>44.8444</td>
<td>13.99</td>
<td>.000</td>
<td>0.21</td>
</tr>
<tr>
<td>EXP x G</td>
<td>1</td>
<td>5.6972</td>
<td>1.78</td>
<td>.188</td>
<td>0.03</td>
</tr>
<tr>
<td>S x EXP / G</td>
<td>54</td>
<td>3.2053</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SNG x EXP</td>
<td>2</td>
<td>0.0230</td>
<td>0.02</td>
<td>.984</td>
<td>0.00</td>
</tr>
<tr>
<td>G x SNG x EXP</td>
<td>2</td>
<td>2.7284</td>
<td>1.91</td>
<td>.153</td>
<td>0.03</td>
</tr>
<tr>
<td>S x SNG x EXP / G</td>
<td>108</td>
<td>1.4267</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>335</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 5: 2 x 3 x 2 Repeated Measures Analysis of Variance of Accuracy Scores by Group (Musicians vs. Nonmusicians).
Figure 3: Mean Accuracy Ratings for Musicians and Nonmusicians Collapsed Across Expressive Condition by Song Performed
between means on Figure 3 reach a level of statistical significance, it is evident that musicians are scoring consistently higher than nonmusicians in terms of musical accuracy. If musicians’ unexpressive version of Amazing Grace is treated as an anomaly or outlier, this trend becomes even more evident and approaches statistical significance ($p < .06$).

**Research Question 3:** Does the number of years participating in a school performing ensemble affect the ability to perform a musical passage with expression?

To examine the effect that the number of years of participation in a performing ensemble has on the ability to perform a musical passage with expression, the musician group was stratified into six groups of differing levels of participation. Each group represented the number of years spent in performance courses. Table 6 presents the means and standard deviations for the ability to perform a musical passage expressively by subjects’ number of years of performing ensemble involvement. As there was only one subject that fell into the three year category, this subject will be eliminated from the discussion as these scores would be expected to regress as the number of subjects in this group increased. Examination of Table 6 reveals that the most expressive performances were found on America the Beautiful, and were evidenced for those participating in performing ensembles for six ($M=6.25$) and seven ($M=5.70$) years. The two groups with the most experience also scored higher than other groups on the expressive performance of Amazing Grace (six years, $M=5.20$; seven years, $M=5.43$). However, this is not the case for the unknown song. Here the group with the least amount of experience scored highest on expressive performances (two years, $M=5.00$).
<table>
<thead>
<tr>
<th>Years</th>
<th>N</th>
<th>Amazing Grace</th>
<th></th>
<th>America the Beautiful</th>
<th></th>
<th>Unknown</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Expressive</td>
<td>Unexpressive</td>
<td>Expressive</td>
<td>Unexpressive</td>
<td>Expressive</td>
</tr>
<tr>
<td>Two</td>
<td>5</td>
<td>4.40</td>
<td>1.50</td>
<td>3.80</td>
<td>3.35</td>
<td>5.00</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2.16</td>
<td>0.50</td>
<td>1.30</td>
<td>1.88</td>
<td>1.87</td>
</tr>
<tr>
<td>Three</td>
<td>1</td>
<td>2.75</td>
<td>2.00</td>
<td>7.00</td>
<td>1.00</td>
<td>2.00</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Four</td>
<td>3</td>
<td>3.50</td>
<td>3.00</td>
<td>4.33</td>
<td>4.00</td>
<td>4.17</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.50</td>
<td>1.00</td>
<td>2.08</td>
<td>2.00</td>
<td>0.88</td>
</tr>
<tr>
<td>Five</td>
<td>4</td>
<td>3.00</td>
<td>2.25</td>
<td>5.38</td>
<td>3.94</td>
<td>4.88</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.83</td>
<td>1.89</td>
<td>1.11</td>
<td>3.00</td>
<td>0.63</td>
</tr>
<tr>
<td>Six</td>
<td>5</td>
<td>5.20</td>
<td>3.10</td>
<td>6.25</td>
<td>2.80</td>
<td>3.75</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.45</td>
<td>1.75</td>
<td>1.03</td>
<td>2.16</td>
<td>1.85</td>
</tr>
<tr>
<td>Seven</td>
<td>10</td>
<td>5.43</td>
<td>3.80</td>
<td>5.70</td>
<td>3.70</td>
<td>4.45</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.42</td>
<td>1.67</td>
<td>1.07</td>
<td>2.21</td>
<td>0.83</td>
</tr>
<tr>
<td>Total</td>
<td>20</td>
<td>4.55</td>
<td>2.89</td>
<td>5.31</td>
<td>3.45</td>
<td>4.37</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.70</td>
<td>1.64</td>
<td>1.45</td>
<td>2.13</td>
<td>1.33</td>
</tr>
</tbody>
</table>

Table 6: Means (M) and Standard Deviations (SD) of Expression Scores for Number of Years of Participation in a School Performing Ensemble as Reported by Song and Expressive Condition.
All groups were successful in conveying a difference between expressive and unexpressive performances. It is important to note that a larger magnitude of differentiation between expressive and unexpressive performances was evidenced for groups with more years of ensemble participation. The only exception to this was the two year groups’ performance of *Amazing Grace*.

To examine overall expressive and unexpressive renderings, means were collapsed across songs (see Table 7). With the exception of the two year group, a linear trend for expressive performances is evidenced. Beginning with the three year group, scores on performances with expression increase as number of years in ensembles increases. Figure 4 graphically represents this trend by depicting the mean ratings of groups by the number of years of participation in a performing ensemble.

Expression ratings were subjected to a three factor (6 x 3 x 2) Repeated Measures Analysis of Variance in order to determine if significant differences existed among the variables in question. Results, presented in Table 8, indicated two sources of statistical significance: (a) main effects for song, $F(2, 44) = 3.95$, $p < .05$; and (b) main effects for expression $F(1, 22) = 15.61$, $p < .001$. Most relevant to the study, the latter finding indicates that when collapsing over groups and songs a significant difference between expressive and unexpressive renderings is evidenced (collapsed $M=4.50$ and collapsed $M=2.95$, respectively, $p < .001$). No significant difference was found for group membership, however ($F(5, 22) = 0.42$, $p > .80$).
<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>Expressive</th>
<th>Unexpressive</th>
</tr>
</thead>
<tbody>
<tr>
<td>Two</td>
<td>5</td>
<td>4.40</td>
<td>2.75</td>
</tr>
<tr>
<td></td>
<td>M</td>
<td>0.49</td>
<td>0.88</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Three</td>
<td>1</td>
<td>3.92</td>
<td>1.67</td>
</tr>
<tr>
<td></td>
<td>M</td>
<td>2.20</td>
<td>0.45</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Four</td>
<td>3</td>
<td>4.00</td>
<td>3.22</td>
</tr>
<tr>
<td></td>
<td>M</td>
<td>0.36</td>
<td>0.57</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Five</td>
<td>4</td>
<td>4.42</td>
<td>3.46</td>
</tr>
<tr>
<td></td>
<td>M</td>
<td>1.02</td>
<td>0.86</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Six</td>
<td>5</td>
<td>5.07</td>
<td>2.90</td>
</tr>
<tr>
<td></td>
<td>M</td>
<td>1.02</td>
<td>0.14</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Seven</td>
<td>10</td>
<td>5.19</td>
<td>3.68</td>
</tr>
<tr>
<td></td>
<td>M</td>
<td>0.53</td>
<td>0.11</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>28</td>
<td>4.50</td>
<td>2.95</td>
</tr>
<tr>
<td></td>
<td>M</td>
<td>0.48</td>
<td>0.65</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 7: Means (M) and Standard Deviations (SD) of Expression Scores by Number of Years of School Ensemble Participation Collapsed Across Songs.
Figure 4: Mean Expression Ratings by Number of Years of Participation in a School Performing Ensemble.
<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>p</th>
<th>Eta Sqr.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Between Subjects</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Years (Y)</td>
<td>5</td>
<td>5.6436</td>
<td>1.28</td>
<td>.306</td>
<td>0.23</td>
</tr>
<tr>
<td>Subjects (S) / Y</td>
<td>22</td>
<td>4.3945</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Within Subjects</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Song (SNG)</td>
<td>2</td>
<td>8.3011</td>
<td>3.95</td>
<td>.026</td>
<td>0.15</td>
</tr>
<tr>
<td>SNG x Y</td>
<td>10</td>
<td>3.6537</td>
<td>1.74</td>
<td>.102</td>
<td>0.28</td>
</tr>
<tr>
<td>S x SNG / Y</td>
<td>44</td>
<td>2.0100</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Expression (EXP)</td>
<td>1</td>
<td>62.5333</td>
<td>15.61</td>
<td>.001</td>
<td>0.42</td>
</tr>
<tr>
<td>EXP x Y</td>
<td>5</td>
<td>1.6877</td>
<td>0.42</td>
<td>.829</td>
<td>0.09</td>
</tr>
<tr>
<td>S x EXP / Y</td>
<td>22</td>
<td>4.0071</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SNG x EXP</td>
<td>2</td>
<td>3.9356</td>
<td>1.70</td>
<td>.195</td>
<td>0.07</td>
</tr>
<tr>
<td>Y x SNG x EXP</td>
<td>10</td>
<td>2.5787</td>
<td>1.11</td>
<td>.375</td>
<td>0.20</td>
</tr>
<tr>
<td>S x SNG x EXP / Y</td>
<td>44</td>
<td>2.3185</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>167</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 8: 6 x 3 x 2 Repeated Measures Analysis of Variance of Expression Scores by Number of Years of Participation in a School Performing Ensemble (Y).
Research Question 4: Does the number of years participating in a school performing ensemble affect the ability to perform a musical passage accurately?

Table 9 presents the means and standard deviations for the ability to perform a musical selection accurately by number of years spent in a school performing ensemble. Again, the group with three years of experience is eliminated from the discussion because of a small sample size (N=1). As in research question two, a trend to perform the expressive version of the songs more accurately than the unexpressive version is evidenced- the only exception being the performance of America the Beautiful by subjects who participated in ensembles for four years (expressive, M=5.67, unexpressive, M=6.67).

Groups with six or seven years experience achieved higher accuracy means on expressive renderings of Amazing Grace (M=5.20 and M=5.15, respectively) and America the Beautiful (M=6.50 and M=5.78, respectively). With the exception of four year subjects’ performance of the unexpressive unknown song (M=2.67) and five year subjects’ expressive version (M=6.44), means for the unknown song tended to stabilize for expressive and unexpressive conditions. In general, accuracy means tended to be rather high across all groups with only three instances of the mean score falling below a rating of 3. Each of these sub-3 ratings occurred on a unexpressive performance, and were evidenced for the two year group on Amazing Grace (M=2.15), the five year group on Amazing Grace (M=2.75), and the four year group on the unknown (M=2.67).
<table>
<thead>
<tr>
<th>Years</th>
<th></th>
<th>Amazing Grace</th>
<th></th>
<th>America the Beautiful</th>
<th></th>
<th>Unknown</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>Expressive</td>
<td>Unexpressive</td>
<td>Expressive</td>
<td>Unexpressive</td>
<td>Expressive</td>
</tr>
<tr>
<td>Two</td>
<td>5</td>
<td>4.75</td>
<td>2.15</td>
<td>4.50</td>
<td>4.00</td>
<td>5.20</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.70</td>
<td>0.86</td>
<td>1.73</td>
<td>2.62</td>
<td>1.64</td>
</tr>
<tr>
<td>Three</td>
<td>1</td>
<td>3.50</td>
<td>2.00</td>
<td>6.00</td>
<td>7.00</td>
<td>2.00</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Four</td>
<td>3</td>
<td>4.25</td>
<td>3.67</td>
<td>5.67</td>
<td>6.67</td>
<td>4.75</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.98</td>
<td>2.08</td>
<td>0.58</td>
<td>0.58</td>
<td>1.56</td>
</tr>
<tr>
<td>Five</td>
<td>4</td>
<td>3.75</td>
<td>2.75</td>
<td>5.44</td>
<td>4.75</td>
<td>6.44</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.71</td>
<td>0.96</td>
<td>1.71</td>
<td>2.63</td>
<td>0.97</td>
</tr>
<tr>
<td>Six</td>
<td>5</td>
<td>5.20</td>
<td>3.45</td>
<td>6.50</td>
<td>4.35</td>
<td>5.35</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.45</td>
<td>1.77</td>
<td>0.71</td>
<td>2.67</td>
<td>1.45</td>
</tr>
<tr>
<td>Seven</td>
<td>10</td>
<td>5.15</td>
<td>4.30</td>
<td>5.78</td>
<td>4.80</td>
<td>4.78</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.99</td>
<td>1.36</td>
<td>0.91</td>
<td>1.48</td>
<td>1.44</td>
</tr>
<tr>
<td>Total</td>
<td>28</td>
<td>4.73</td>
<td>3.39</td>
<td>5.63</td>
<td>4.85</td>
<td>5.09</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.32</td>
<td>1.53</td>
<td>1.24</td>
<td>2.08</td>
<td>1.54</td>
</tr>
</tbody>
</table>

Table 9: Means (M) and Standard Deviations (SD) of Accuracy Scores by Number of Years Participation in a School Performing Ensemble as Reported by Song and Expressive Condition.
Accuracy ratings were subjected to a three way (6 x 3 x 2), Type III Repeated Measures Analysis of Variance to investigate differences. The results, presented in Table 10, indicated three sources of statistical significance: (a) main effects for song, $F(2, 44) = 16.72$, $p < .001$; (b) main effects for expression $F(1, 22) = 4.48$, $p < .05$; and (c) first order interaction effects between years of participation and song, $F(10, 44) = 3.02$, $p < .01$.

Figure 5 depicts the first order interaction between years of ensemble participation and song collapsed across expresssive condition. To explicate further the nature of the differences between the means, Tukey's HSD post hoc follow up procedures were employed using the harmonic mean to control for unequal cell Ns (described in Kennedy and Bush, 1985). Significant differences were evidenced between four year subjects' performance of Amazing Grace and America the Beautiful (Tukey $q = 3.61$, $p < .05$, $df = 44$), and the same subjects' performance of America the Beautiful and the unknown song (Tukey $q = 4.08$, $p < .05$, $df = 44$). Other than this, no other significant differences were found.

In regard to significant main effects, accuracy on performances of America the Beautiful (collapsed $M = 5.45$) was significantly higher than for Amazing Grace (collapsed $M = 3.74$) and the unknown song (collapsed $M = 4.35$) when collapsed across all groups and expressive conditions ($p < .01$). Additionally, expressive performances were performed significantly more accurate when collapsed across groups and songs (collapsed $M = 4.94$ to collapsed $M = 4.09$, respectively, $p < .01$).
<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>p</th>
<th>Eta Sqr.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Between Subjects</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Years (Y)</td>
<td>5</td>
<td>3.0791</td>
<td>0.47</td>
<td>.797</td>
<td>0.10</td>
</tr>
<tr>
<td>Subjects (S) / Y</td>
<td>22</td>
<td>6.5963</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Within Subjects</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Song (SNG)</td>
<td>2</td>
<td>26.0684</td>
<td>16.72</td>
<td>.000</td>
<td>0.43</td>
</tr>
<tr>
<td>SNG x Y</td>
<td>10</td>
<td>4.6932</td>
<td>3.02</td>
<td>.005</td>
<td>0.41</td>
</tr>
<tr>
<td>S x SNG / Y</td>
<td>44</td>
<td>1.5560</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Expression (EXP)</td>
<td>1</td>
<td>19.0766</td>
<td>4.48</td>
<td>.046</td>
<td>0.17</td>
</tr>
<tr>
<td>EXP x Y</td>
<td>5</td>
<td>1.3865</td>
<td>0.33</td>
<td>.892</td>
<td>0.07</td>
</tr>
<tr>
<td>S x EXP / Y</td>
<td>22</td>
<td>4.2547</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SNG x EXP</td>
<td>2</td>
<td>2.1560</td>
<td>1.64</td>
<td>.207</td>
<td>0.07</td>
</tr>
<tr>
<td>Y x SNG x EXP</td>
<td>10</td>
<td>1.8790</td>
<td>1.43</td>
<td>.201</td>
<td>0.25</td>
</tr>
<tr>
<td>S x SNG x EXP / Y</td>
<td>44</td>
<td>1.3185</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>167</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 10: 6 x 3 x 2 Repeated Measures Analysis of Variance of Accuracy Scores by Number of Years of Participation in a School Performing Ensemble (Y).
Figure 5: Mean Accuracy Ratings by Number of Years of Participation in a School Performing Ensemble.
Presentation of the Data from Phase II

Results pertaining to research questions five through eight are presented below. These data are intended to elucidate the nature of the school performing ensemble experience as it relates to the ability to perceive expression in music and the ability to perceive how accurately a musical passage is performed. Recalling that research question #1 found that musicians performed with significantly more differentiation between expressive and unexpressive performances than nonmusicians, it was concluded that these significant changes in expressivity by musicians evidenced from Phase I would prove more salient for listeners attempting to perceive performers’ expressive intent. Therefore, it was decided to analyze only the ratings of the musicians’ performances from Phase I in Phase II of the study. It was assumed that if subjects were able to perceive changes in expressive intent, their best chance for success would be in using this sample. As in Phase I, it was decided to treat expression and accuracy data as separate analyses for clarity.

Research Question 5: Is there a significant difference between subjects who have had school performing ensemble experience and subjects who have not their ratings of perceived expression in a musical passage?

Descriptive statistics indicating means and standard deviations for subjects’ perceived expression are presented in Table 11. With the exception of musicians’ ratings for Amazing Grace (expressive, $M=4.23$; unexpressive, $M=4.47$), all subjects rated expressive performances as more expressive than
<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td><strong>Amazing Grace</strong></td>
<td><strong>America the Beautiful</strong></td>
<td><strong>Unknown</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Expressive</strong></td>
<td><strong>Unexpressive</strong></td>
<td><strong>Expressive</strong></td>
<td><strong>Unexpressive</strong></td>
</tr>
<tr>
<td>Musicians</td>
<td>73</td>
<td>4.23</td>
<td>4.47</td>
<td>4.82</td>
<td>3.50</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.00</td>
<td>1.44</td>
<td>0.89</td>
<td>1.43</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4.07</td>
<td>3.62</td>
<td>0.90</td>
<td>1.07</td>
</tr>
<tr>
<td>Nonmusicians</td>
<td>56</td>
<td>4.28</td>
<td>4.25</td>
<td>4.42</td>
<td>3.96</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.01</td>
<td>1.44</td>
<td>1.00</td>
<td>1.36</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4.09</td>
<td>3.99</td>
<td>0.84</td>
<td>1.17</td>
</tr>
<tr>
<td>Total</td>
<td>129</td>
<td>4.25</td>
<td>4.37</td>
<td>4.65</td>
<td>3.70</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.00</td>
<td>1.44</td>
<td>0.96</td>
<td>1.41</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4.08</td>
<td>3.78</td>
<td>0.87</td>
<td>1.12</td>
</tr>
</tbody>
</table>

Table 11: Means (M) and Standard Deviations (SD) of Musicians' and Nonmusicians' Ratings of Perceived Expression as Reported by Song and Expressive Condition.
unexpressive performances. Musicians and nonmusicians rated the expressive version of America the Beautiful as being the most expressive ($M=4.82$ and $M=4.42$, respectively), and the unexpressive version of this song as least expressive ($M=3.50$ and $M=3.96$, respectively). Not surprisingly, America the Beautiful is evidenced as being most salient in musicians' and nonmusicians' discrimination between expressive and unexpressive performances. Musicians ($\Delta = 1.32$) were able to demonstrate this discrimination to a larger degree than nonmusicians ($\Delta = 0.46$), however. Furthermore, musicians demonstrated greater discrimination between expressive and unexpressive performance than did nonmusicians regardless of song, although in the case of Amazing Grace, this was in an opposite direction of the intended expressive condition.

Perceived expression ratings were subjected to a three way (2 x 3 x 2), Type III Repeated Measures Analysis of Variance to investigate differences. Because assumptions of homogeneity of covariance were violated (Mauchly's $W = 41.61, p < .01$), the implementation of the Geisser-Greenhouse conservative $F$ test was employed to protect against spurious findings. The results, presented in Table 12, indicated five sources of statistical significance: (a) main effects for song, $F(1.6, 196.9) = 8.42, p < .001$; (b) main effects for expressive condition, $F(1, 127) = 21.46, p < .001$; (c) first order interaction effects between group membership and expressive condition, $F(1, 127) = 4.10, p < .05$; (d) first order interaction effects between song and expressive condition, $F(1.6, 198.2) = 15.38, p < .001$; and, (e) second order interaction effects between groups, song, and expressive condition, $F(1.6, 198.2) = 4.87, p < .05$. 

102
<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>p</th>
<th>η² Sqr.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Between Subjects</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Groups (G)</td>
<td>1.0</td>
<td>0.4826</td>
<td>0.18</td>
<td>.671</td>
<td>0.00</td>
</tr>
<tr>
<td>Subjects (S) / G</td>
<td>127.0</td>
<td>6.5963</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Within Subjects</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Song (SNG)</td>
<td>1.6</td>
<td>11.1794</td>
<td>8.42</td>
<td>.000</td>
<td>0.06</td>
</tr>
<tr>
<td>SNG x G</td>
<td>1.6</td>
<td>1.5916</td>
<td>1.20</td>
<td>.303</td>
<td>0.01</td>
</tr>
<tr>
<td>S x SNG / G</td>
<td>196.9</td>
<td>1.3278</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Expression (EXP)</td>
<td>1.0</td>
<td>24.0892</td>
<td>21.46</td>
<td>.000</td>
<td>0.15</td>
</tr>
<tr>
<td>EXP x G</td>
<td>1.0</td>
<td>4.5975</td>
<td>4.10</td>
<td>.045</td>
<td>0.03</td>
</tr>
<tr>
<td>S x EXP / G</td>
<td>127.0</td>
<td>1.1222</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SNG x EXP</td>
<td>1.6</td>
<td>20.5071</td>
<td>15.38</td>
<td>.000</td>
<td>0.11</td>
</tr>
<tr>
<td>G x SNG x EXP</td>
<td>1.6</td>
<td>6.4953</td>
<td>4.87</td>
<td>.015</td>
<td>0.04</td>
</tr>
<tr>
<td>S x SNG x EXP / G</td>
<td>198.2</td>
<td>1.3342</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td>658.5</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Note.** Statistics reflect the Geisser-Greenhouse Conservative F test as assumptions of homogeneity of covariance were violated.

Table 12: 2 x 3 x 2 Repeated Measures Analysis of Variance of Ratings of Perceived Expression by Group (Musicians vs. Nonmusicians).
Figures 6 displays a view of the second order interaction. It is apparent when viewing this graph that both group membership and expressive condition influenced perceived expression, but this was moderated by song. In particular, *Amazing Grace* performances were heard in the opposite direction of the intended expressive condition by musicians (expressive $M=4.23$, unexpressive, $M=4.47$), whereas nonmusicians indicated perceiving little change in expressive ($M=4.28$) and unexpressive ($M=4.25$) performances. Nonmusicians continued this trend of perceiving little change across expressive conditions in performances of *America the Beautiful* (expressive, $M=4.42$; unexpressive, $M=3.96$) and the unknown (expressive, $M=4.09$; unexpressive, $M=3.99$), while musicians indicated larger differences in the intended direction of expressivity for both *America the Beautiful* (expressive, $M=4.82$; unexpressive, $M=3.50$) and the unknown (expressive, $M=4.07$, unexpressive, $M=3.62$). To determine if any significant differences existed among these 6 pairs of means, post hoc Scheffe follow up procedures were employed. Using this conservative follow up measure, only one significant difference was discovered: musicians' ratings of expression on the expressive version of *America the Beautiful* were significantly higher than their ratings of the unexpressive version ($F(s) = 3.63, p < .05, df = 198$).

The first order interaction between groups and expressive condition compared means collapsed over songs. These collapsed means were as follows: musicians, expressive condition, $M=4.37$; musicians, unexpressive condition, $M=3.86$; nonmusicians, expressive condition, $M=4.27$; nonmusicians unexpressive condition, $M=4.07$. Post hoc procedures revealed musicians' ratings of expressive performances to be significantly higher than their ratings of unexpressive performances ($F(s) = 3.06, p < .05, df = 127$).
Figure 6: Mean Ratings of Perceived Expression Depicting Second Order Interaction Between Group, Song, and Expressive Condition
As in prior analyses, main effects for expression collapsed over groups and song was significant with the expressive condition receiving higher ratings of expression ($p < .001$). Also, a significant main effect for song when collapsed over group and expressive condition was evidenced with the unknown song being rated significantly lower overall ($p < .05$).

Research Question 6: Is there a significant difference between subjects who have had school performing ensemble experience and subjects who have not in their ratings of perceived accuracy in a musical passage?

Means and standard deviation for subjects' perception of musical accuracy are presented in Table 13. In general, musicians tended to use higher ratings than nonmusicians in evaluating musical accuracy. Both musicians and nonmusicians rated the *America the Beautiful* expressive condition as being the most accurate (musicians, $M=5.02$; nonmusicians, $M=4.71$).

Interestingly, *America the Beautiful* also received the lowest accuracy rating from musicians and nonmusicians on the unexpressive condition (musicians, $M=3.70$; nonmusicians, $M=3.42$). The tendency to rate the expressive performance as more accurate than the unexpressive performance was evidenced in both groups for the unknown song as well (musicians $M=4.46$ to $M=3.89$; nonmusicians $M=4.30$ to $M=3.54$). However, this did not hold true for *Amazing Grace* where both musicians and nonmusicians rated the unexpressive version as more accurate (musicians $M=4.96$ to $M=4.61$; nonmusicians $M=4.54$ to $M=4.44$).
<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>Amazing Grace</th>
<th>America the Beautiful</th>
<th>Unknown</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Expressive</td>
<td>Unexpressive</td>
<td>Expressive</td>
</tr>
<tr>
<td>Musicians</td>
<td>73</td>
<td>4.61</td>
<td>4.96</td>
<td>5.02</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.99</td>
<td>1.37</td>
<td>0.98</td>
</tr>
<tr>
<td>Nonmusicians</td>
<td>56</td>
<td>4.44</td>
<td>4.54</td>
<td>4.61</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.93</td>
<td>1.28</td>
<td>0.93</td>
</tr>
<tr>
<td>Total</td>
<td>129</td>
<td>4.54</td>
<td>4.78</td>
<td>4.84</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.96</td>
<td>1.34</td>
<td>0.98</td>
</tr>
</tbody>
</table>

Table 13: Means (M) and Standard Deviations (SD) of Musicians’ and Nonmusicians’ Ratings of Perceived Accuracy as Reported by Song and Expressive Condition.
Accuracy means were subjected to a three way (2 x 3 x 2), Type III Repeated Measures Analysis of Variance to investigate differences. Because assumptions of homogeneity of covariance were violated (Mauchly's W = 35.59, p < .01), the implementation of the Geisser-Greenhouse conservative F test was employed to protect against spurious findings. The results, presented in Table 14, indicated four sources of statistical significance: (a) main effects for group, F(1, 127) = 4.60, p < .05; (b) main effects for song, F(1.7, 219.5) = 34.59, p < .001; (c) main effects for expression, F(1, 127) = 65.85, p < .001; and (d) first order interaction effects between song and expressive condition, F(1.6, 203.8) = 52.21, p < .001.

Figure 7 displays a graphic representation of cell means collapsed across groups for the first order interaction between song and expressive condition. To explicate further the differences between the collapsed means, Scheffe post hoc follow up procedures were employed to reveal any significant differences. While these tests revealed no significant differences (p > .10) between the expressive (M=4.53) and unexpressive (M=4.75) performances of Amazing Grace, significant differences were evidenced between the expressive and unexpressive performances of America the Beautiful (F(s) = 17.55, p < .01, df = 203), and the expressive and unexpressive performances of the unknown song (F(s) = 4.86, p < .05, df = 203). In each of the latter instances, expressive performances were regarded as significantly more accurate. Furthermore, the unexpressive performance of Amazing Grace received a significantly higher accuracy rating than America the Beautiful (F(s) = 15.67, p < .01, df = 203) and the unknown (F(s) = 11.31, p < .01, df = 208). These higher accuracy ratings on Amazing Grace account for the significant main effect for song found in the
<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>p</th>
<th>Eta Sqr.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Between Subjects</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Groups (G)</td>
<td>1.0</td>
<td>16.9333</td>
<td>4.60</td>
<td>.034</td>
<td>0.04</td>
</tr>
<tr>
<td>Subjects (S) / G</td>
<td>127.0</td>
<td>3.6835</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Within Subjects</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Song (SNG)</td>
<td>1.7</td>
<td>27.5301</td>
<td>34.59</td>
<td>.000</td>
<td>0.21</td>
</tr>
<tr>
<td>SNG x G</td>
<td>1.7</td>
<td>0.1485</td>
<td>0.19</td>
<td>.830</td>
<td>0.00</td>
</tr>
<tr>
<td>S x SNG / G</td>
<td>219.5</td>
<td>0.7961</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Expression (EXP)</td>
<td>1.0</td>
<td>60.3384</td>
<td>65.85</td>
<td>.000</td>
<td>0.34</td>
</tr>
<tr>
<td>EXP x G</td>
<td>1.0</td>
<td>0.5176</td>
<td>0.57</td>
<td>.454</td>
<td>0.00</td>
</tr>
<tr>
<td>S x EXP / G</td>
<td>127.0</td>
<td>0.9166</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SNG x EXP</td>
<td>1.6</td>
<td>43.4957</td>
<td>52.21</td>
<td>.000</td>
<td>0.29</td>
</tr>
<tr>
<td>G x SNG x EXP</td>
<td>1.6</td>
<td>0.8742</td>
<td>1.05</td>
<td>.339</td>
<td>0.01</td>
</tr>
<tr>
<td>S x SNG x EXP / G</td>
<td>203.8</td>
<td>0.8333</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>686.9</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Note.** Statistics reflect the Geisser-Greenhouse Conservative F test as assumptions of homogeneity of covariance were violated.

Table 14: 2 x 3 x 2 Repeated Measures Analysis of Variance of Ratings of Perceived Accuracy by Group (Musicians vs. Nonmusicians).
Figure 7: Mean Ratings of Perceived Accuracy Depicting Interaction Between Song and Expressive Condition.
Analysis of Variance table. That is, overall accuracy ratings on *Amazing Grace* were significantly higher (p < .001) than *America the Beautiful* and the unknown song.

A main effect was evidenced for expressive condition when collapsing over groups and songs, indicating that accuracy ratings were significantly higher (p < .001) on expressive performances (collapsed M=4.57) than on unexpressive performances (collapsed M=4.01). Furthermore, in examining main effects for group differences on ratings of musical accuracy, the musician group gave significantly higher ratings (p < .05) than the nonmusician group overall (M=4.40 to M=4.14, respectively).

**Research Question 7:** Does the number of years participating in a school performing ensemble affect ratings of perceive expression in musical passage?

Means and standard deviations for perceived expression ratings are presented in Table 15 by number of years of participation in a school ensemble. It was decided to eliminate the nine year group from discussion because of a small sample size (N=1). For *Amazing Grace*, the majority of groups rated the unexpressive performances as being more expressive than the expressive performances. The only exceptions to this were the four year group (expressive M=4.17; unexpressive, M=4.00) and the eight year group (expressive, M=4.69; unexpressive, M=4.00). For *America the Beautiful*, the majority of groups perceived the expressive condition as being more expressive than the unexpressive. The only exception was the three year group that rated both conditions about the same (expressive, M=4.75; unexpressive, M=4.80). It is important to note that the largest differences
<table>
<thead>
<tr>
<th>Years</th>
<th>N</th>
<th>Amazing Grace</th>
<th></th>
<th>America the Beautiful</th>
<th></th>
<th>Unknown</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Expressive</td>
<td>Unexpressive</td>
<td>Expressive</td>
<td>Unexpressive</td>
<td>Expressive</td>
<td>Unexpressive</td>
</tr>
<tr>
<td>One</td>
<td>11</td>
<td>4.10</td>
<td>4.27</td>
<td>4.75</td>
<td>3.45</td>
<td>4.03</td>
<td>3.48</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.80</td>
<td>1.74</td>
<td>1.05</td>
<td>1.19</td>
<td>0.97</td>
<td>1.27</td>
</tr>
<tr>
<td>Two</td>
<td>11</td>
<td>3.94</td>
<td>4.64</td>
<td>3.43</td>
<td>3.33</td>
<td>4.05</td>
<td>3.59</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.02</td>
<td>1.69</td>
<td>0.99</td>
<td>1.54</td>
<td>0.96</td>
<td>1.09</td>
</tr>
<tr>
<td>Three</td>
<td>5</td>
<td>4.20</td>
<td>4.60</td>
<td>4.75</td>
<td>4.80</td>
<td>4.27</td>
<td>3.35</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.94</td>
<td>1.95</td>
<td>0.92</td>
<td>1.02</td>
<td>1.07</td>
<td>0.48</td>
</tr>
<tr>
<td>Four</td>
<td>6</td>
<td>4.17</td>
<td>4.00</td>
<td>5.17</td>
<td>4.81</td>
<td>3.54</td>
<td>4.17</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.57</td>
<td>1.26</td>
<td>0.85</td>
<td>1.02</td>
<td>1.09</td>
<td>1.55</td>
</tr>
<tr>
<td>Five</td>
<td>5</td>
<td>3.98</td>
<td>4.40</td>
<td>5.13</td>
<td>2.90</td>
<td>4.48</td>
<td>3.55</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.10</td>
<td>1.14</td>
<td>0.69</td>
<td>1.52</td>
<td>0.52</td>
<td>1.37</td>
</tr>
<tr>
<td>Six</td>
<td>8</td>
<td>4.31</td>
<td>4.63</td>
<td>5.34</td>
<td>3.13</td>
<td>3.75</td>
<td>3.46</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.12</td>
<td>1.41</td>
<td>0.58</td>
<td>1.28</td>
<td>1.02</td>
<td>1.07</td>
</tr>
<tr>
<td>Seven</td>
<td>23</td>
<td>4.44</td>
<td>4.65</td>
<td>4.76</td>
<td>3.20</td>
<td>4.25</td>
<td>3.72</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.92</td>
<td>1.77</td>
<td>0.80</td>
<td>1.27</td>
<td>0.81</td>
<td>0.98</td>
</tr>
<tr>
<td>Eight</td>
<td>3</td>
<td>4.69</td>
<td>4.00</td>
<td>4.33</td>
<td>3.22</td>
<td>3.60</td>
<td>3.24</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.24</td>
<td>1.00</td>
<td>1.04</td>
<td>1.84</td>
<td>0.85</td>
<td>3.75</td>
</tr>
<tr>
<td>Nine</td>
<td>1</td>
<td>3.50</td>
<td>3.00</td>
<td>5.00</td>
<td>3.33</td>
<td>4.67</td>
<td>3.75</td>
</tr>
<tr>
<td>Total</td>
<td>73</td>
<td>4.23</td>
<td>4.47</td>
<td>4.82</td>
<td>3.50</td>
<td>4.07</td>
<td>3.62</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.00</td>
<td>1.44</td>
<td>0.89</td>
<td>1.43</td>
<td>0.90</td>
<td>1.07</td>
</tr>
</tbody>
</table>

Table 15: Means (M) and Standard Deviations (SD) of Ratings of Perceived Expression by Number of Years Participation in a School Performing Ensemble as Reported by Song and Expressive Condition.
between expressive and unexpressive performances were evidenced for *America the Beautiful* as well. Similar to *America the Beautiful*, perceived expression for the unknown was in the intended direction for the majority of groups. The four year group was the only exception to this, rating the unexpressive condition (M = 4.14) higher than the expressive (M = 3.54).

Expression means were subjected to a three way (9 x 3 x 2), Type III Repeated Measures Analysis of Variance to investigate differences. Because assumptions of homogeneity of covariance were violated (Mauchly's W = 0.3306, p < .01), the implementation of the Geisser-Greenhouse conservative F test was employed to protect against spurious findings. The results, presented in Table 16, indicated three sources of statistical significance: (a) main effects for song, F(1.5, 97.5) = 3.47, p < .05; (b) main effects for expression, F(1, 64) = 16.75, p < .001; and (c) first order interaction effects between song and expressive condition, F(1.4, 90.9) = 7.20, p < .01.

Figure 8 depicts the first order interaction between song and expressive condition collapsed across groups. As discussed above, perceived expression is in the intended direction for *America the Beautiful* and the unknown song with the largest discrimination evidenced for *America the Beautiful*, whereas *Amazing Grace* mean expression ratings are in the opposite direction of the intended expressive performance. Scheffe post hoc follow up procedures found significant differences between the expressive (M = 4.85) and unexpressive (M = 3.58) rendition of *America the Beautiful* (F(s) = 17.59, p < .01, df = 90.9). Furthermore, the unexpressive performance of *Amazing Grace* was perceived as being significantly more expressive than the unexpressive performances of *America the Beautiful* and the unknown song (p < .01).
<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>p</th>
<th>Eta Sgr.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Between Subjects</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Years (Y)</td>
<td>8.0</td>
<td>0.8466</td>
<td>0.31</td>
<td>.960</td>
<td>0.04</td>
</tr>
<tr>
<td>Subjects (S) / Y</td>
<td>64.0</td>
<td>2.7472</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Within Subjects</strong></td>
<td>287.9</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Song (SNG)</td>
<td>1.5</td>
<td>4.5882</td>
<td>3.47</td>
<td>.047</td>
<td>0.05</td>
</tr>
<tr>
<td>SNG x Y</td>
<td>12.2</td>
<td>1.5712</td>
<td>1.19</td>
<td>.303</td>
<td>0.13</td>
</tr>
<tr>
<td>S x SNG / Y</td>
<td>97.5</td>
<td>1.3245</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Expression (EXP)</td>
<td>1.0</td>
<td>16.6106</td>
<td>16.75</td>
<td>.000</td>
<td>0.21</td>
</tr>
<tr>
<td>EXP x Y</td>
<td>8.0</td>
<td>0.9035</td>
<td>0.91</td>
<td>.514</td>
<td>0.10</td>
</tr>
<tr>
<td>S x EXP / Y</td>
<td>64.0</td>
<td>0.9924</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SNG x EXP</td>
<td>1.4</td>
<td>11.8736</td>
<td>7.20</td>
<td>.004</td>
<td>0.10</td>
</tr>
<tr>
<td>Y x SNG x EXP</td>
<td>11.4</td>
<td>1.2889</td>
<td>0.78</td>
<td>.662</td>
<td>0.09</td>
</tr>
<tr>
<td>S x SNG x EXP / Y</td>
<td>90.9</td>
<td>1.6502</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>359.9</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Note.** Statistics reflect the Geisser-Greenhouse Conservative F test as assumptions of homogeneity of covariance were violated.

Table 16: 9 x 3 x 2 Repeated Measures Analysis of Variance of Ratings of Perceived Expression by Number of Years in a School Performing Ensemble.
Figure 8: Mean Ratings of Perceived Expression Across All Years of School Ensemble Participation Depicting Interaction Between Song and Expressive Condition.
The most salient finding of the analysis, a significant main effect for expression ($p < .001$), revealed that all groups rated expressive performances ($M = 4.36$) as significantly more expressive than unexpressive performances ($M = 3.80$) despite the Amazing Grace anomaly. Number of years in a performing ensemble did not have a significant effect on perceptions of expressivity ($p > .50$).

**Research Question 8:** Does the number of years participating in a school performing ensemble affect ratings of perceived accuracy in a musical passage?

Means and standard deviations for perceived accuracy ratings are presented in Table 17 by number of years of participation in a school ensemble. As in the previous analysis, it was decided to eliminate the nine year group from discussion because of its small sample size ($N = 1$). For Amazing Grace, the majority of groups rated the unexpressive performances as being more accurate than the expressive performances. The only exceptions to this were the one year group (expressive $M = 4.33$; unexpressive, $M = 4.18$), the four year group (expressive, $M = 4.95$; unexpressive, $M = 4.83$) and the eight year group (expressive, $M = 4.89$; unexpressive, $M = 4.67$). For America the Beautiful, all groups perceived the expressive condition as being more accurate than the unexpressive. In addition, all groups gave the expressive version of America the Beautiful their highest rating of accuracy. Accuracy ratings for the unknown song were higher for expressive performances for all of the groups as well. In general, most of the accuracy scores remain stable across expressive condition, usually varying .6 points or less.
<table>
<thead>
<tr>
<th>Years</th>
<th>N</th>
<th>Amazing Grace</th>
<th></th>
<th>America the Beautiful</th>
<th></th>
<th>Unknown</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Expressive</td>
<td>Unexpressive</td>
<td>Expressive</td>
<td>Unexpressive</td>
<td>Expressive</td>
<td>Unexpressive</td>
</tr>
<tr>
<td>One</td>
<td>11</td>
<td>4.33</td>
<td>4.18</td>
<td>4.52</td>
<td>3.41</td>
<td>4.24</td>
<td>3.43</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.78</td>
<td>1.83</td>
<td>0.99</td>
<td>1.23</td>
<td>0.83</td>
<td>1.26</td>
</tr>
<tr>
<td>Two</td>
<td>11</td>
<td>4.62</td>
<td>5.27</td>
<td>4.75</td>
<td>3.33</td>
<td>4.41</td>
<td>3.93</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.04</td>
<td>1.10</td>
<td>0.70</td>
<td>1.55</td>
<td>0.85</td>
<td>1.30</td>
</tr>
<tr>
<td>Three</td>
<td>5</td>
<td>4.20</td>
<td>5.00</td>
<td>4.80</td>
<td>4.40</td>
<td>4.40</td>
<td>4.00</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.45</td>
<td>1.00</td>
<td>0.33</td>
<td>0.86</td>
<td>0.82</td>
<td>1.08</td>
</tr>
<tr>
<td>Four</td>
<td>6</td>
<td>4.95</td>
<td>4.83</td>
<td>5.58</td>
<td>4.47</td>
<td>4.49</td>
<td>4.38</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.64</td>
<td>1.72</td>
<td>1.28</td>
<td>1.49</td>
<td>1.26</td>
<td>1.31</td>
</tr>
<tr>
<td>Five</td>
<td>5</td>
<td>3.82</td>
<td>4.60</td>
<td>4.62</td>
<td>3.20</td>
<td>4.23</td>
<td>3.70</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.70</td>
<td>1.14</td>
<td>0.75</td>
<td>0.89</td>
<td>0.80</td>
<td>1.60</td>
</tr>
<tr>
<td>Six</td>
<td>8</td>
<td>4.97</td>
<td>5.25</td>
<td>5.47</td>
<td>4.17</td>
<td>4.48</td>
<td>4.47</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.21</td>
<td>1.75</td>
<td>1.13</td>
<td>1.07</td>
<td>0.64</td>
<td>1.11</td>
</tr>
<tr>
<td>Seven</td>
<td>23</td>
<td>4.75</td>
<td>5.21</td>
<td>5.21</td>
<td>3.58</td>
<td>4.61</td>
<td>3.75</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.84</td>
<td>1.17</td>
<td>0.93</td>
<td>1.43</td>
<td>0.83</td>
<td>1.10</td>
</tr>
<tr>
<td>Eight</td>
<td>3</td>
<td>4.89</td>
<td>4.67</td>
<td>4.92</td>
<td>3.61</td>
<td>4.59</td>
<td>4.08</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.45</td>
<td>1.15</td>
<td>1.81</td>
<td>2.26</td>
<td>1.52</td>
<td>1.77</td>
</tr>
<tr>
<td>Nine</td>
<td>1</td>
<td>4.25</td>
<td>5.00</td>
<td>5.50</td>
<td>4.67</td>
<td>4.83</td>
<td>4.25</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>73</td>
<td>4.60</td>
<td>4.96</td>
<td>5.02</td>
<td>3.70</td>
<td>4.46</td>
<td>3.89</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.99</td>
<td>1.37</td>
<td>0.98</td>
<td>1.36</td>
<td>0.84</td>
<td>1.20</td>
</tr>
</tbody>
</table>

Table 17: Means (M) and Standard Deviations (SD) of Ratings of Perceived Accuracy by Number of Years Participation in a School Performing Ensemble as Reported by Song and Expressive Condition.
Accuracy means were subjected to a three way (9 x 3 x 2), Type III Repeated Measures Analysis of Variance to investigate differences. Because assumptions of homogeneity of covariance were violated (Mauchly's $W = 13.15$, $p < .01$), the implementation of the Geisser-Greenhouse conservative $F$ test was employed to protect against spurious findings. The results, presented in Table 18, indicated three sources of statistical significance: (a) main effects for song, $F(1.8, 114.4) = 5.49$, $p < .01$; (b) main effects for expression, $F(1, 64) = 11.18$, $p < .01$; and (c) first order interaction effects between song and expressive condition, $F(1.7, 107.7) = 14.89$, $p < .001$.

Figure 9 depicts the first order interaction between song and expressive condition collapsed across groups. Scheffe post hoc follow up procedures found significant differences between the expressive ($M=5.04$) and unexpressive ($M=3.87$) rendition of *America the Beautiful* ($F(s) = 34.95$, $p < .001$, $df = 107.7$) as well as between the expressive ($M=4.48$) and unexpressive ($M=4.00$) performances of the unknown ($F(s) = 4.91$, $p < .05$, $df = 107.7$). Furthermore, significant differences were found between the unexpressive performance of *Amazing Grace* and the unexpressive performances of *America the Beautiful* and the unknown song ($p < .01$) with *Amazing Grace* being perceived as more accurate.

A significant main effect for expression was found ($p < .01$) with expressive performances ($M=4.68$) being rated significantly higher in accuracy than unexpressive performances ($M=4.25$) across groups and songs.

Regarding differences in ratings of accuracy as determined by the number of years of participation in a performing ensemble, no significant differences were found between the groups ($p > .50$).
<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>p</th>
<th>Eta Sqr.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Between Subjects</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Years (Y)</td>
<td>8</td>
<td>3.6111</td>
<td>0.85</td>
<td>.565</td>
<td>0.10</td>
</tr>
<tr>
<td>Subjects (S) / Y</td>
<td>64</td>
<td>4.2592</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Within Subjects</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Song (SNG)</td>
<td>1.8</td>
<td>4.5297</td>
<td>5.49</td>
<td>.007</td>
<td>0.08</td>
</tr>
<tr>
<td>SNG x Y</td>
<td>14.3</td>
<td>0.6233</td>
<td>0.75</td>
<td>.718</td>
<td>0.09</td>
</tr>
<tr>
<td>S x SNG / Y</td>
<td>114.4</td>
<td>0.8253</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Expression (EXP)</td>
<td>1.0</td>
<td>9.9147</td>
<td>11.18</td>
<td>.001</td>
<td>0.15</td>
</tr>
<tr>
<td>EXP x Y</td>
<td>8</td>
<td>0.5413</td>
<td>0.61</td>
<td>.766</td>
<td>0.07</td>
</tr>
<tr>
<td>S x EXP / Y</td>
<td>64</td>
<td>0.8875</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SNG x EXP</td>
<td>1.7</td>
<td>12.5643</td>
<td>14.89</td>
<td>.000</td>
<td>0.19</td>
</tr>
<tr>
<td>Y x SNG x EXP</td>
<td>13.5</td>
<td>0.5222</td>
<td>0.62</td>
<td>.839</td>
<td>0.07</td>
</tr>
<tr>
<td>S x SNG x EXP / Y</td>
<td>107.7</td>
<td>0.8444</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td>398.4</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Note.** Statistics reflect the Geisser-Greenhouse Conservative F test as assumptions of homogeneity of covariance were violated.

Table 18: 9 x 3 x 2 Repeated Measures Analysis of Variance of Ratings of Perceived Accuracy by Number of Years in a School Performing Ensemble.
Figure 9: Mean Ratings of Perceived Accuracy Across All Years of School Ensemble Participation Depicting Interaction Between Song and Expressive Condition.
CHAPTER 5

SUMMARY, DISCUSSION, CONCLUSIONS, AND RECOMMENDATIONS FOR FURTHER RESEARCH

Summary

The purpose of this study was to examine the effects of school performing ensemble participation on the ability to perceive expression in music and to perform a musical passage expressively. In order to examine perception and performance, the study was divided into two phases. Phase I addressed the performance of musical expression using a minimal skill musical device, while Phase II examined the perception of musical expression. Because expression is contingent to a large degree on the accuracy of musical performance, it was decided to include an assessment of musical accuracy in each phase of the study as well.

A survey of literature revealed that many educational theorists, philosophers, and music theorists believe that performance and perception are mutually informative dimensions. Studies suggest that musicians convey knowledge and understanding of underlying musical structure through expressive performance. Expressive performance, it would appear, is guided by perceptual understanding. Furthermore, many educational theorists advocate the study of performance to enhance perceptive abilities.
While many studies were found that examined these relationships with expert musicians, relatively few studies were found that examined the impact of performing on the generative and receptive abilities of younger students and novices. Even less were discovered that investigated the impact of school performing ensemble experiences on these abilities.

In order to assess the effect of the performing ensemble on generative and receptive abilities, the present investigation compared collegiate level subjects who had prior experience in school performing ensemble courses only (musicians) with those who had not (nonmusicians). In Phase I, an experiment was devised in which musicians (N=28) and nonmusicians (N=28) used a minimal skill musical device to perform two familiar songs and one unfamiliar song under two conditions: performing expressively and unexpressively. Final performances were recorded and sent to four expert judges for evaluation. Judges rated performances in terms of musical expression and accuracy. In Phase II, an experiment was devised in which musicians (N=73) and nonmusicians (N=56) rated a random sample of the performances generated in Phase I of the study. A total of 54 performances from Phase I were evaluated by the subjects for both expression and accuracy.

Reliability and internal consistency measures for Phase I and Phase II were calculated. A moderate level of inter-rater reliability was found for the judges in Phase I for ratings of expression (r_{s(avg)} = .57) and accuracy (r_{s(avg)} = .69).¹ Judges proved to be highly internally consistent in ratings of both

¹. Judges were not trained in order to establish a similar frame of reference, thus, to some degree assessments varied in the extent of usage of the rating scale. Judges appear to have agreed consistently in the direction of their ratings, but varied in the range, degree, and standard employed in assessment.
expression ($r_s = .84$) and accuracy ($r_s = .88$). Subjects in Phase II were found to exhibit a much lower rate of internal consistency than the judges on ratings of expression ($r_s = .35$) and accuracy ($r_s = .36$). Stratifying this group into musicians and nonmusicians revealed that musicians were more internally consistent than nonmusicians in ratings of expression ($r_s = .39$ to $r_s = .29$, respectively) and accuracy ($r_s = .44$ to $r_s = .20$, respectively).

Because no significant effects were found for order, gender, or type of performing ensemble experience (i.e., band, orchestra, or choir) in either Phase I or Phase II (all $p > .30$), these variables were eliminated from subsequent analyses. The results of the experiment, reported in response to the research questions, are presented below. In order to answer questions pertaining to the effect of years of participation, the musician group was stratified into groups that corresponded to the subjects’ number of years of experience in school performing ensembles. In Phase I these groups consisted of two, three, four, five, six, and seven years of participation; in Phase II, the groups were one, two, three, four, five, six, seven, eight, and nine years of participation.

Question 1: Is there a significant difference between subjects who have had school performing ensemble experience and subjects who have not in the ability to perform a musical passage with expression?

A significant first order interaction ($p < .05$) was evidenced between group (musicians and nonmusicians) and expressive condition. Subsequent post hoc analysis revealed that musicians’ expressive performances were rated significantly more expressive than their unexpressive performances ($p < .05$), while no significant difference was found between nonmusicians’ expressive and unexpressive performances ($p > .10$). Furthermore, post hoc
comparisons revealed that musicians were more expressive on expressive performances ($p < .10$) than nonmusicians. However, no significant difference was found between musicians' and nonmusicians' unexpressive performances ($p > .10$).

**Question 2:** Is there a significant difference between subjects who have had school performing ensemble experience and subjects who have not in the ability to perform a musical passage accurately?

A difference approaching statistical significance was evidenced between the groups on the ability to perform a musical passage accurately ($p < .10$). Specifically, musicians ($M=4.65$) were found to be more accurate in their performances than nonmusicians ($M=4.18$) overall. All subjects as a whole performed expressive renditions of songs significantly more accurately than unexpressive versions ($p < .001$), and were able to perform *America the Beautiful* significantly more accurately than *Amazing Grace* and the unknown song ($p < .001$).

**Question 3:** Does the number of years participating in a school performing ensemble affect the ability to perform a musical passage with expression?

No significant difference ($p > .80$) was found when investigating the effect of the number of years of ensemble participation on the ability to perform a musical passage expressively. However, a linear trend was found that indicated expression ratings on the expressive condition increased as the number of years of participation increased.
Question 4: Does the number of years participating in a school performing ensemble affect the ability to perform a musical passage accurately?

A significant first order interaction was found between the number of years in a performing ensemble and song ($p < .01$). Graphing this interaction revealed that the group with four years of experience was able to perform America the Beautiful significantly more accurately than Amazing Grace and the unknown ($p < .05$). More importantly, the graph illustrated the tendency for those with more experience to be more stable in accuracy across all three songs. No significant main effect was found for differences between groups overall ($p > .75$).

Question 5: Is there a significant difference between subjects who have had school performing ensemble experience and subjects who have not in their ratings of perceived expression in a musical passage?

A significant second order interaction was found between groups, song, and expressive condition ($p < .05$). Musicians were found to perceive Amazing Grace performances in the opposite direction of the intended expressive condition. Specifically, musicians rated the unexpressive version of Amazing Grace ($M=4.47$) as more expressive than the expressive version ($M=4.23$). Post hoc follow up procedures indicated that this was not a significant difference however ($p > .10$). All other expressive conditions on each song were in the intended direction for both groups with the only significant difference evidenced for musicians in distinguishing expressive performances of America the Beautiful from unexpressive performances ($p < .05$).
When collapsing across songs, a significant first order interaction was discovered ($p < .05$). Here, post hoc procedures revealed musicians’ ratings of expressive performances to be significantly higher than their ratings of the unexpressive performances ($p < .05$) despite the *Amazing Grace* anomaly. No significant difference was found between nonmusicians’ ratings of expressive and unexpressive performances ($p > .10$).

Question 6. Is there a significant difference between subjects who have had school performing ensemble experience and subjects who have not in their ratings of perceived accuracy in a musical passage?

A significant difference in overall accuracy ratings was found between musicians and nonmusicians ($p < .05$). Musicians ($M=4.40$) rated performances significantly higher in regard to musical accuracy than nonmusicians ($M=4.14$).

Although not as relevant to the particular question examined, this analysis also revealed a significant first order effect between song and expressive condition ($p < .001$). Post hoc analysis revealed that the expressive versions of *America the Beautiful* and the unknown song were rated significantly more accurate than their unexpressive counterparts ($p < .05$). Furthermore, the unexpressive version of *Amazing Grace* was rated significantly more accurate than the unexpressive version of *America the Beautiful* and the unknown ($p < .05$).
Question 7: Does the number of years participating in a school performing ensemble affect ratings of perceived expression in a musical passage?

No significant difference in the ability to perceive musical expression was found to be attributable to the number of years of participation in a performing ensemble ($p > .50$). However, a significant first order interaction was found between song and expressive condition ($p < .01$). Post hoc analysis revealed that the expressive version of *America the Beautiful* was perceived as significantly more expressive than its unexpressive version for all subjects ($p < .05$). Additionally, the unexpressive version of *Amazing Grace* was perceived as significantly more expressive than the unexpressive versions of *America the Beautiful* and the unknown song ($p < .05$).

Question 8: Does the number of years participating in a school performing ensemble affect ratings of perceived accuracy in a musical passage?

No significant difference in the perception of musical accuracy was found to be attributable to the number of years of participation in a performing ensemble ($p > .50$). However, a significant interaction was found between song and expressive condition. Post hoc analysis revealed that the expressive versions of *America the Beautiful* and the unknown song were perceived as being significantly more accurate than their unexpressive counterparts ($p < .05$). Furthermore, the unexpressive version of *Amazing Grace* was perceived as significantly more accurate than the unexpressive versions of *America the Beautiful* and the unknown song ($p < .01$).
Discussion

The results from Phase I of the study indicate that subjects with school performing ensemble experience are able to perform more expressively and more accurately than subjects without this background. This finding parallels that of other researchers who have demonstrated that the communication of musical expression is contingent on the experience and training of the performer (Palmer, 1996b; Sloboda, 1985). Also similar to Palmer’s (1996b) findings, the ability of the musicians in this study to perform with a larger differentiation between expressive and unexpressive performances than the nonmusicians reflects an ability to be more fluid in the application of expressive parameters.

It is important to note, however, that all subjects were able to communicate some difference between expressive and unexpressive performance regardless of group membership, although musicians were able to demonstrate this difference to a greater degree than nonmusicians. This suggests that general acculturation may have some impact on the development of expressive performance. Lending support to this theory is the effect size evidenced in question #1. Based on the eta square statistic, it is evident that the contribution of the expressive condition variable accounts for the greatest amount of variance (35%) while the contribution of group membership by expressive condition is much less substantial (7%). However, this is qualified somewhat by the most salient finding in question #1: the ability of the musician group to demonstrate the most difference between expressive and unexpressive performances.
Nonmusicians can be seen attempting to communicate expressive performances, but do not succeed in producing a significant difference between expressive and unexpressive performances. This may be because nonmusicians lack the specific musical understanding to communicate these expressive features. In light of Lerdahl and Jackendoff's (1983) work, this would seem to suggest that the cognizance of underlying musical structure is essential to the effective communication of musical expression. It could be that musicians, because of their prior training, have an advantage over nonmusicians in that they are sensitized to underlying phrase and harmonic structures, and, thus, are able to affect salient expressive performances as a consequence. Nonetheless, the ability of nonmusicians to portray expressive and unexpressive performances in consistently different ways is reflective of studies that indicate the similarities of musicians and nonmusicians and the ability of these students to respond to music in similar ways (Crist, 1996; Nakamura, 1987; Madsen, Byrnes, Capperella-Sheldon, & Brittin, 1993; Madsen & Fredrickson, 1993; Madsen, Byrnes, Capperella-Sheldon, & Brittin, 1993). Future research might examine the specific performances of both groups to see if expressive nuance is tied to an understanding of musical structure in some way, and if this understanding is the same or different for musicians and nonmusicians.

In examining the performance of the individual songs, several interesting patterns emerge. Of the familiar songs, the easiest in terms of rhythmic and passing tone demands, America the Beautiful, was performed most expressively and accurately by both musicians and nonmusicians. In fact, musicians and nonmusicians scored similarly in terms of musical accuracy for both the expressive (M=5.63 to M=5.09, respectively) and
unexpressive (M=4.85 to M=4.45, respectively) versions. Even with similar accuracy, musicians were still able to engender a greater magnitude of difference (Δ = 1.86) between unexpressive and expressive performances than nonmusicians (Δ = .89). Turning to Amazing Grace, a more difficult song in terms of rhythm and passing tones, nonmusicians seem to have struggled much more than musicians in reproducing the tune accurately (nonmusicians, M=3.86; musicians, M=4.73). Expressive scores are also quite low for nonmusicians (M=3.54). Viewing the data in this light, it could be that the majority of the nonmusicians efforts went into learning to perform the song accurately, and, perhaps as a consequence, they simply did not have time within the 10 minute limit imposed to develop an expressive performance.

What could be indicated in the data, then, is an ability of musicians to quickly assimilate new musical material. Viewing the data for the unknown song reinforces this theory. Although not statistically significant (p > .10), musicians (M= 5.09) were observed to be more accurate in their performance of the expressive version than nonmusicians (M=4.31). It could be because of this ability to learn the unknown song quickly that musicians were able to produce a more expressive performance than nonmusicians (Δ = .88). Again, this is reflective of Palmer’s (1996b) work with music majors and professional musicians in which musicians with more experience are evidenced as being able to learn new musical material more quickly, and, therefore, are able to render expressive performances sooner. Viewed another way, the data could be regarded in terms of musical memory. It may be that musicians’ musical memory is developed beyond that of nonmusicians, and they simply
remember "how the song goes" with less listenings. Future research might investigate the musical memory of this population to clarify this as a potential factor.

Extending somewhat from the idea of musical memory, the ability of musicians to learn the pieces more quickly may have afforded more practice time on expressive aspects of the music. As the study did not include an investigation of the use of the 10 minute time limit, it is difficult to conclude if this is the case. It would seem logical to assume that if accuracy is mastered quickly, then more effort can be made on expressive features in subsequent practice trials. Furthermore, as musicians are supposedly trained in the use of practice to improve performance, a difference would be expected between the groups in terms of how the practice time was used. A more thorough analysis of the use of practice time in future studies may reveal different practicing techniques at work.

Lastly, it seems apparent that accuracy affected ratings of musical expression to a large degree. In rating expression, judges may have been reacting to the accuracy of musical performance and not necessarily to expressive features. Simply put, it could be that judges were responding to the accuracy of performances first, and, thus, the more accurate a song was performed, the more expressive it was judged to be. This speaks to the confounding nature musical accuracy has on musical expression. Without correct notes and rhythms, any performance would seem inferior expressively. Summarizing Sloboda (1994): performers cannot vary the expressive nuance of any particular note or notes at random; "it matters which notes we vary and how" (p. 154). Therefore, if expressive nuance did occur in an inaccurate performance, it is unlikely that the manipulation of
expressive parameters would correspond to the listener's sense of underlying musical structure. Hence, the inaccurate performance would be deemed unexpressive because of the lack of fit between structural properties and expressive nuance. Essentially, varying expressive parameters on wrong notes and rhythms may give the listener the impression that expressive nuance is simply being manipulated "randomly." The old adage that the mastery of notes and rhythms comes before expressive interpretation may indeed have substantive meaning.

Turning to the effect of the number of years of performing ensemble participation on performance abilities, no significant difference in the ability to perform with expression was evidenced. However, a trend was found that illustrated expression ratings improving as the number of years of participation in a performing ensemble increased. This trend is consistent with other performance studies that have investigated the effect of experience on expressive performance (Palmer, 1989, 1996b; Sloboda, 1983). Additionally, data from the accuracy ratings did evidence a significant interaction between number of years of participation and song (p < .05). Realizing that America the Beautiful was the easier of the familiar performances, it is understandable why groups with lesser experience improved significantly on this performance (p < .05). However, subjects with more experience were found to be more stable in their ability to render a performance accurately across both of the familiar songs and the unfamiliar song. As discussed above, this may be because extensive experience in music improves one's ability to reproduce new musical stimuli more quickly and with fewer listenings. An interesting anomaly did arise in the groups with two and five years of experience, however. These groups performed the unknown song more
accurately than either of the two familiar songs. It is difficult to speculate as to why this occurred. Perhaps, the difference may be attributable to a preference for the unknown song, or because the subjects were not familiar with this piece, they practiced the song longer and performed it more accurately as a consequence. Future studies investigating how students with varying levels of experience practice familiar and unfamiliar pieces could help clarify this finding.

The results from Phase II of the study indicate that subjects with school performing ensemble experience perceive expression and accuracy in musical performance differently than subjects without this background. Specifically, musicians displayed a significant difference between their ratings of expressive and unexpressive performances overall (\(p < .05\)), whereas, nonmusicians did not display a significant difference in their expressive and unexpressive perceptions (\(p > .10\)). Furthermore, musicians and nonmusicians displayed a significant difference in their rating of musical accuracy (\(p < .05\)) with musicians rating the accuracy of performances higher overall. As in Phase I these findings are qualified by effect size. Based on the eta square statistic, it is the expressive condition variable that accounts for the most variance (15%) in perceived expression ratings, whereas group membership by expressive condition is responsible for much less of the variance (3%). Likewise, in perceived accuracy ratings, it is the expressive condition (34%) and song (21%) which are accounting for the majority of the variance, while group is much less substantial (4%).

In regard to the perception of expression, it is important to note that the above finding was moderated by song, as a significant (\(p < .05\)) second order interaction was evidenced between group, song, and expressive
condition. Specifically, musicians rated the unexpressive version of Amazing Grace as more expressive than the expressive version (M=4.47 to M=4.23, respectively). This was different than the nonmusicians who indicated perceiving both versions of Amazing Grace as having relatively the same amount of expressivity (expressive, M=4.28; unexpressive, M=4.25).

Examining this finding in relation to accuracy ratings indicates that musicians and nonmusicians rated the unexpressive version of Amazing Grace as more accurate than the expressive version (musicians, M=4.96 to M=4.61, respectively; nonmusicians, M=4.54 to M=4.44, respectively). Thus, musicians' ratings of musical expression may have been confounded by their perceptions of accuracy. It is interesting to note that although nonmusicians also rated the unexpressive version of Amazing Grace as more accurate, this did not seem to impact their ratings of expressivity to as great a degree.

Perhaps nonmusicians are not as influenced by correct rhythms when rating musical expression as musicians are, or perhaps as Madsen and Geringer (1990), Geringer and Madsen (1996), and Johnson (1996a) have found, musicians and nonmusicians attend to different musical parameters and apply different standards of evaluation when listening to music.

Other than the Amazing Grace anomaly, both groups were able to perceive each performance in the intended expressive direction. This supports other research studies (Crist, 1996; Kendall & Carterette, 1990) that have found musicians and nonmusicians to be similar in their ability to label performances as expressive or unexpressive. However, the magnitude of the difference evidenced in musicians' ability to distinguish expressive from unexpressive performance indicates that musicians are more likely to be able to make confident assessments of expressive performance as compared to
nonmusicians. Thus, rating performances for expression and accuracy may involve more refined skills of musicianship than simple categorization tasks. This would seem quite significant when regarding a musical device such as expression, because of its Gestalt nature. As this study and others have found (Crist, 1996; Johnson, 1996b; Kendall & Carterette, 1990), expression is easily identified regardless of musical background. However, making appropriate judgments as to the appropriateness or degree of musical expression seems to be a more advanced task that is predicated on musical background/ experience. Had Amazing Grace been eliminated from the present study or another, “easier” piece used in its stead, the magnitude of difference between musicians’ and nonmusicians’ discrimination skills might have proved even greater.

The ability of musicians in this study to make significant distinctions between expressive and unexpressive performances could be a result of focus of attention as well. It is evident, from the internal consistency coefficients generated, that musicians were more consistent in their evaluations of performances than nonmusicians in terms of expression (.39 to .29, respectively) and accuracy (.44 to .20, respectively). When comparing this to the expert judges from Phase I an interesting trend is evidenced. Experts are most consistent (expression, .84; accuracy, .88), followed by musicians, and then nonmusicians. This could be an indication of two possibilities working either in isolation or in concert: one, musical training plays a significant role in the development of musical standards by which to judge musical performance consistently; and/or, two, musicians are able to focus on evaluative tasks for extended amounts of time better than nonmusicians. Subsequent studies may investigate these hypotheses for further clarification.
It is also interesting to note that novice and expert musicians both had a higher internal consistency for accuracy ratings, whereas expression ratings were not as consistent. Although this is a small difference, it may be that accuracy is somewhat easier to rate than expression as there may be a consistent standard employed. Nonmusicians demonstrated an opposite trend, rating expression more consistently than accuracy. Again, this finding, although not directly studied by the investigation, may lend support to studies that have found that musicians and nonmusicians typically focus on different parameters when listening to music (Geringer & Madsen, 1996; Madsen & Geringer, 1990).

Finally, in regard to musicians and nonmusicians ratings of musical expression and accuracy, a qualification must be made. As Phase II examined only the performances of the musicians from Phase I, it is necessary to compare each groups' expression and accuracy ratings with that of the experts from Phase I to determine which group is rating accuracy in a manner consistent with the experts. Examining expression ratings reveals that experts in Phase I rated musicians' expressive performances ($M=4.74$) much higher than their unexpressive performances ($M=3.22$). In Phase II, musician listeners were most similar to this, rating expressive performances ($M=4.37$) higher than unexpressive performances ($M=3.86$). Nonmusicians showed little differentiation between the two expressive conditions, rating expressive performances ($M=4.27$) slightly higher than unexpressive performances ($M=4.07$). Evident here is the trend for musicians with greater experience to be more differentiating in their assessment of a performance's expressive merit.
Comparing the expert judges' overall ratings of the musician performers' accuracy from Phase I (M=4.65) to musician listeners (M=4.40) and nonmusician listeners (M=4.14) in Phase II reveals that the musicians in Phase II were more similar to the experts in terms of assessing musical accuracy as well. As discussed above, this may indicate a standard of accuracy that is learned through the study of music.

The final research questions pertaining to the effect of the number of years of participation in a performing ensemble on perceptive abilities revealed no significant differences between groups. The only significant differences found for expression and accuracy related to an interaction between song and expressive condition (p < .01). Here again, the interaction was caused by Amazing Grace on both ratings of expression and accuracy. Furthermore, the most salient perceived expressive difference was evidenced in America the Beautiful which also received the largest perceived difference in accuracy ratings across expressive conditions. These findings continue to reinforce that musicians' ratings of expression and accuracy are highly interdependent. Simplifying somewhat, performances must be perceived as accurate in order to be perceived as expressive.

In conclusion, the results of this study suggest that the performance and perception of musical expression are positively influenced by school performing ensemble participation. Although all subjects exhibited ability in both generative and receptive activities, these abilities were more pronounced in those subjects with experience in performing ensembles. This finding, in itself, suggests that perception and production are mutually informative. By engaging in performance, subsequent perceptive abilities are positively influenced. In regard to the effect of the length of time spent in an
ensemble on productive and perceptive abilities, no significant differences were discerned in this study. However, a trend for those with more experience to score higher and render performances more consistently accurate on the performance task was evidenced. These data suggest that those with more experience have the ability to assimilate new musical material more quickly, and/or the ability to be more adaptive to new musical situations in comparison to those with lesser experience. Further research in this area is warranted to clarify these apparent trends.

In regard to the findings of this study, one further qualification should be made. As the sample consisted of volunteer subjects who had elected to take a music course freely as a general elective in their collegiate course work, it could be argued that these subjects were predisposed to music. In other words, although the study carefully considered each subject's musical background in regard to performance, it did not take into account subjects' listening habits, musical preferences, parents' value of music, etc. It could be that students who choose to take music courses as general electives value music to a greater degree than those who do not elect to take these courses. Furthermore, the listening tendencies and knowledge of different musical genres of these students might be quite extensive and diverse in comparison to students who elect not to take these courses. As a consequence, this may be why the differences between the musician and nonmusician groups employed in the present investigation are somewhat small. Had another population been used for nonmusicians (e.g., one that did not elect to take music courses as an elective at the collegiate level), the differences between groups might have been more pronounced. Future research could assist in determining the validity of this hypothesis.
Conclusions

The following conclusions are drawn from the analysis of the data:

1. School performing ensemble participation has a significant positive effect on the performance of musical expression. Although those with performing ensemble participation can engender a greater distinction between expressive and unexpressive performance, those without this experience can demonstrate a subtle difference between expressive and unexpressive performance.

2. School performing ensemble participation has a positive, but not statistically significant, effect on the ability to perform a musical passage accurately.

3. A trend would indicate that the number of years of participation in a performing ensemble has a positive effect on the performance of musical expression. Although not found to be significant, the trend suggests that as performing ensemble experience increases, so does the ability to communicate an expressive performance.

4. A trend would indicate that the number of years of performance in a performing ensemble is associated positively with the ability to perform a musical passage accurately. Although not found to be significant, the trend would indicate that those with more experience are more consistent in the ability to perform a selection accurately.

5. School performing ensemble participation appears to have a positive effect on perceptive abilities. Those with performing ensemble experience are able to discriminate between expressive and unexpressive
performances to a greater extent than those who lack this background. Those without performing ensemble experience are able to perceive differences between expressive and unexpressive performances, however, but not to as great a degree.

6. Those with performing ensemble experience rate musical accuracy differently than those without this experience. Those with performing ensemble experience resemble musical experts in their ratings of musical accuracy more so than those without this experience.

7. The number of years of performing ensemble experience does not appear to effect perceptions of musical expression.

8. The number of years of performing ensemble experience does not appear to effect perceptions of musical accuracy.

9. Musical expression is contingent to a large degree on the accuracy of musical performance. Musicians’ ratings of musical expression seem to be influenced by accuracy to a greater degree than those of nonmusicians.

10. Musicians are more internally consistent in evaluative tasks than nonmusicians, and are more consistent in terms of accuracy ratings than expression ratings. On the other hand, nonmusicians are more consistent in regard to expression ratings than accuracy ratings, suggesting that musicians and nonmusicians may be attending to different features when listening to music.
Recommendations for Further Research

1. Further research into the differences between those with school musical experience and those without is warranted to appreciate more fully the learning that is generated in school music curricula.

2. Research that addresses musical memory between these populations should seek to clarify the effect of experience on musicians’ and nonmusicians’ ability to contend with new musical situations. By understanding this, it would be possible to glean an understanding of how these groups partition cognitive resources in listening to, evaluating and responding to music.

3. An investigation of musicians’ and nonmusicians’ focus of attention during musical listening could help clarify the apparent discrepancy evidenced in internal consistency ratings found in this study. Understanding this phenomenon would afford a better view of how these groups attend to music.

4. Continued effort to tease out the effect of musical accuracy on perceptions of musical expression is needed. By understanding this relationship, the pedagogy of musical expression can be informed and improved.

5. As an anecdotal aside, it is important to note that the use of the Instant Pleasure software seemed to be an effective way to equalizing the skill necessary in order to render a musical performance. Furthermore, many subjects expressed that using the device was quite enjoyable and indicated that they would like to obtain a copy of the software in order to perform for friends and family. As mentioned earlier, all subjects reported that their
performances using *Instant Pleasure* were indicative of their intentions. Future studies should continue to employ this device as it allows nonmusicians to communicate their perceptions of music through a performer's perspective. Additionally, the use of this device as an educative medium could be quite effective in allowing nonmusicians the opportunity to think and act as performers.
BIBLIOGRAPHY


Nierman, G. E. (1983). The role of private music instruction in the
development of high school students' ability to describe musical
events. *Bulletin of the Council for Research in Music Education*, 73,
15-27.

Oakley, D. L. (1972). The cumulative attainment by Missouri high school
seniors of the music learnings stated in the music curriculum guides
published by the Missouri State Department of Education.
Unpublished doctoral dissertation, Indiana University, Bloomington,
Indiana.


*Journal of Experimental Psychology: Human Perception and
Performance, 15* (3), 331-346.


*Music Perception, 14* (1), 23-56.

prosody. *Journal of Memory and Language.*

melodic discrimination test and the background variables of
prospective music students. *Bulletin of the Council for Research in


Radocy, R. E., & Boyle, J. D. (1997). *Psychological foundations of musical

experiences on preschool children’s melodic perception. *Journal of
Research in Music Education, 31* (2), 133-145.

Prentice Hall.


APPENDIX A

MUSICAL EXCERPTS
Amazing Grace

Amazing Grace Composite Rhythm
America the Beautiful

America the Beautiful Composite Rhythm
APPENDIX B

MUSICAL BACKGROUND SURVEY
Musical Background Survey

The following questionnaire is designed to gather information about your musical background, reasons for taking Music 141, and this course’s influence on your attitudes about music. Please respond as accurately as possible.

All responses will be held in the strictest confidence. Your name will not be associated with this survey in any way.

Participant # ________________

1. Gender (circle one): Male Female

2. Class Rank (circle one): Freshman Sophomore Junior Senior Other ______

3. What is your major at OSU? __________________________

4. What was your main reason(s) for taking this course?

5. Do you currently play a musical instrument?
   yes no If NO, skip to question 6
   If yes, what instrument(s): ____________________________
   How long have you played? ______

6. if you do not currently play an instrument, have you ever played a musical instrument?
   yes no If NO, skip to question 7
   If yes, what instrument(s): ____________________________
   How long did you play? ______

7. Do you regularly participate in a musical group(s) at the present time?
   yes no If NO, skip to question 9
8. Please indicate the groups in which you participate currently and how long you have participated: (check all that apply)

| ______ | Orchestra | How long? |
| ______ | Concert Band | How long? |
| ______ | Marching Band | How long? |
| ______ | Jazz Ensemble | How long? |
| ______ | Rock Ensemble | How long? |
| ______ | Folk/Ethnic Music Ensemble | How long? |
| ______ | Church Choir | How long? |
| ______ | Community or Civic Choir | How long? |
| ______ | Musical Theater (Broadway type or Operatic) | How long? |
| ______ | Other | Please list: |

9. Have you ever participated in a musical group(s) in the past? [If you answered yes to question 7, this applies to group(s) in which you are not currently participating.]

yes no If NO, skip to question 11

10. Please indicate the groups in which you have participated in the past and the length of your involvement: (check all that apply)

| ______ | Orchestra | How long did you participate? |
| ______ | Concert Band | How long did you participate? |
| ______ | Marching Band | How long did you participate? |
| ______ | Jazz Ensemble | How long did you participate? |
| ______ | Rock Ensemble | How long did you participate? |
| ______ | Folk/Ethnic Music Ensemble | How long did you participate? |
| ______ | Church Choir | How long did you participate? |
| ______ | Community or Civic Choir | How long did you participate? |
| ______ | Musical Theater | How long did you participate? |
| ______ | Other | Please list: |

11. Have you ever taken private lessons on an instrument or voice?

yes no If NO, skip to question 12

If yes, what instrument(s)?

How long did you take lessons?

12. Would you consider yourself self-taught (i.e., no lessons or school participation) on any instrument?

yes no If NO, skip to question 13

If yes, what instrument(s)?

How long have you played?

Do you still play? yes no
13. Growing up, did you have a piano/keyboard in the home?    yes    no

14. Please circle the number of years enrolled in High School (Grades 9-12):

<table>
<thead>
<tr>
<th></th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Band</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Orchestra</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>Chorus</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Music Appreciation</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>Music Theory</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
</tbody>
</table>

What instrument(s)?

15. Please circle the number of years enrolled in Middle/Junior High School (Grades 6-8):

<table>
<thead>
<tr>
<th></th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Band</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Orchestra</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Chorus</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Music Appreciation</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Music Theory</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
</tbody>
</table>

What instrument(s)?

16. How many years has it been since you participated in a performing ensemble?    

(If you still participate indicate by answering “still participate”)
APPENDIX C

POST EXPERIMENT FOLLOW-UP SURVEY
Post Experiment Follow-Up Survey

Participant # ________________

Are you enrolled in Mus 140 or Mus 141? ________________

What day and time does your regular class meet? ________________

What day and time does your recitation meet? ________________

Gender (circle one): Male Female

Class Rank (circle one): Freshman Sophomore Junior Senior Other ________________

1. Do you feel that your recordings are representative of your intentions? Yes No

   If No, why not? ________________

2. Did you find the software easy to use? Yes No

3. Did you recognize the Unfamiliar Song? Yes No

   If Yes, what do you think the title is? ________________

5. What type of strategies did you use to complete the task?

6. What do you think the differences are between expressive/musical performance and unmusical/mechanical performance?

7. When you signed up for this experiment, did you sign up as a: (circle one)
   person with high school performing experience only
   person with no prior musical training/experience

Thank you for participating in this experiment!!

165
APPENDIX D

LIKERT SCALE RESPONSE FORM
Likert Scale Response Form
(sample page)

Please evaluate the following performances on a scale of 1 to 7 in regard to:

(a) musical accuracy- 1 = poor to 7 = excellent; and

(b) expression/musicality- 1 = unexpressive/without expression to 7 = highly expressive.

<table>
<thead>
<tr>
<th>#</th>
<th>Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>musical accuracy [poor]</td>
</tr>
<tr>
<td></td>
<td>expression [unexpressive]</td>
</tr>
<tr>
<td>2.</td>
<td>musical accuracy [poor]</td>
</tr>
<tr>
<td></td>
<td>expression [unexpressive]</td>
</tr>
<tr>
<td>3.</td>
<td>musical accuracy [poor]</td>
</tr>
<tr>
<td></td>
<td>expression [unexpressive]</td>
</tr>
<tr>
<td>4.</td>
<td>musical accuracy [poor]</td>
</tr>
<tr>
<td></td>
<td>expression [unexpressive]</td>
</tr>
<tr>
<td>5.</td>
<td>musical accuracy [poor]</td>
</tr>
<tr>
<td></td>
<td>expression [unexpressive]</td>
</tr>
<tr>
<td>6.</td>
<td>musical accuracy [poor]</td>
</tr>
<tr>
<td></td>
<td>expression [unexpressive]</td>
</tr>
<tr>
<td>7.</td>
<td>musical accuracy [poor]</td>
</tr>
<tr>
<td></td>
<td>expression [unexpressive]</td>
</tr>
<tr>
<td>8.</td>
<td>musical accuracy [poor]</td>
</tr>
<tr>
<td></td>
<td>expression [unexpressive]</td>
</tr>
<tr>
<td>9.</td>
<td>musical accuracy [poor]</td>
</tr>
<tr>
<td></td>
<td>expression [unexpressive]</td>
</tr>
<tr>
<td>10.</td>
<td>musical accuracy [poor]</td>
</tr>
<tr>
<td></td>
<td>expression [unexpressive]</td>
</tr>
</tbody>
</table>